DOCUMENT RESUME

ED 149 329 CS 003 928

AUTHOR	Massaro, Dominic W.
TITLE	Reading and Listening. Technical Report No. 423.
INSTITUTION	Wisconsin Univ., Madison. Research and Development Center for Cognitive Learning.
SPONS AGENCY	National Inst. of Education (DHEW), Washington, D.C.
PUB DATE	Dec 77
CONTRACT	NE-C-00-30065
NOTE	60p.; Report from the Project on Studies in Reading, Language, and Communication
EDRS PRICE	MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS	Abstract Reasoning; *Aural Stimuli; *Cognitive Processes; Communication (Thought Transfer); *Decoding (Reading); Language Research; *Listening Skills; *Models; *Reading Processes; Reading Skills

ABSTRACT

This study develops an information model based on the processes involved in reading and listening as a step toward constructing a single framework for understanding both these skills. Assuming that analogous structures and processes occur in reading and listening, similar experimental questions are asked in both areas. The model incorporates data and theory from a variety of approaches to reading and speech perception. The properties of storage structures and the dynamics of the functional processes are considered, along with the active utilization of various sources of knowledge available to the language receiver. (MAI)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED149329

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Technical Report No. 423

READING AND LISTENING

by

Dominic W. Massaro

Report from the Project on
Studies in Reading, Language, and Communication

Dominic W. Massaro
Faculty Associate

Wisconsin Research and Development
Center for Cognitive Learning
The University of Wisconsin
Madison, Wisconsin

December 1977

CS003928

Published by the Wisconsin Research and Development Center for Cognitive Learning, supported in part as a research and development center by funds from the National Institute of Education, Department of Health, Education, and Welfare. The opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by that agency should be inferred.

Center Contract No. NE-C-00-3-0065

MISSION STATEMENT

The mission of the Wisconsin Research and Development Center is to improve the quality of education by addressing the full range of issues and problems related to individualized schooling. Teaching, learning, and the problems of individualization are given concurrent attention in the Center's efforts to discover processes and develop strategies and materials for use in the schools. The Center pursues its mission by

- conducting and synthesizing research to clarify the processes of school-age children's learning and development
- conducting and synthesizing research to clarify effective approaches to teaching students basic skills and concepts
- developing and demonstrating improved instructional strategies, processes, and materials for students, teachers, and school administrators
- providing assistance to educators which helps transfer the outcomes of research and development to improved practice in local schools and teacher education institutions

The Wisconsin Research and Development Center is supported with funds from the National Institute of Education and the University of Wisconsin.

WISCONSIN RESEARCH AND DEVELOPMENT
CENTER FOR COGNITIVE LEARNING

ACKNOWLEDGEMENT

The author is grateful for the positive contributions of Michael M. Cohen, David Klitzke, Gregg G. Oden, Glen A. Taylor, and Richard L. Venezky.

TABLE OF CONTENTS

		Page
	Acknowledgments	iv
	List of Tables	vii
	List of Figures	ix
	Abstract	xi
I.	Introduction	1
	A Stage Model of Language Processing	2
	Criticisms of Stage Models	3
	Parallel Interactive Models	6
	Summary	7
II.	Feature Detection	9
	Audible Features	9
	Visible Features	11
	Visible Features in Words	13
III.	Primary Recognition	15
	Perceptual Units in Speech	15
	Target Search in Listening	15
	Perceptual Units in Reading	16
	Target Search in Reading	17
	The Word Advantage.	18
	Single Letters vs. Word Units (an Experimental Test).	19
	A Second Critical Test	25
	Descriptions of Orthographic Structure	30
	Semantic vs. Orthographic Context	32
	Controlling for Orthographic Redundancy	34
	Other Evidence	34
IV.	Secondary Recognition	37
	Perceptual and Contextual Contributions to Listening	37
	Phonological Mediation	43
	Visual, Phonological, and Semantic Codes in Reading	47

TABLE OF CONTENTS (CONT.)

		Page
V.	Rehersal and Recoding	51
	Speech Codes in Generated Abstract Memory	51
VI.	Conclusion	55
	References	57

LIST OF TABLES

Table		Page
1	Contexts for the c-e Test Letter as a Function of the Orthographic Context and Serial Position.	27

LIST OF FIGURES

Figure		Page
1	Flow Diagram of the Temporal Course of Reading and Listening.	4
2	The Same Visual Information can be Interpreted Differently in Different Semantic/ Syntactic Contexts	5
3	A Matrix of Ambiguous Characters Having One or More Functional Attributes in Transition Between Letter Pairs (after Naus & Shillman, 1976).	12
4	Predicted and Observed d'values for the Word, Nonword, and Single Letter Conditions as a Function of SOA.	23
5	Predicted and Observed Probabilities of an e Response as a Function of Orthographic Context and Stimulus Value	29
6	Predicted and Observed Percentage of Fluent Restorations as a Function of the Amount of Feature Change, the Syllable, and the Contextual Constraint of the Mispronunciation (after Marslen-Wilson and Welsh, in press).	40
7	Two Stage Models of the Role of Phonological Mediation in Reading	43

I

INTRODUCTION

A paper on Reading and Listening must be motivated by the belief that these two fascinating skills are not unrelated. And, in fact, it has been commonly believed by many that reading is somehow a skill parasitic to listening: that is, all that is necessary in order to read is to translate the written word into its spoken counterpart, and then to listen to the message. Was it St. Augustine's disciple who was amazed to see that his mentor read without moving his lips? Since most literates learn to listen long before they learn to read, it seems natural that reading instruction should exploit what the child knows about listening. Faced with an unknown word in written text, the child is encouraged to decode it (i.e., sound it out) on the assumption that the spoken rendering will be more likely to be recognized. The primacy of speech has been one motivating influence on theories of reading that view reading as a process of going from print to some speech code on the route to the meaning of the message (Gough, 1972). One could also argue that language learning recapitulates the development of spoken and written language since it is a common belief that man spoke and listened long before he wrote and read. However, Wrolstad (1976) in a stimulating and fascinating manifesto, argues just the reverse.

This author rejects the idea that speech is primary and that reading is somehow subservient or dependent on listening to speech. Recent developments in the understanding of sign language (especially American Sign Language, ASL) are illuminating in this regard. It may be the case that some aspects of ASL are acquired earlier than spoken language. An infant raised in an environment where both English and ASL are used will understand some functions in ASL before the analogous functions in English. This finding topples the idea of the primacy of speech as a form of communication. There is also no evidence that speech perception is dependent on reading. However, I will argue that reading and listening can be viewed as exactly analogous processes, the goal of which is to derive (or create) the meaning of a message.

One argument for these being analogous processes can be formulated in terms of convergent evolution: two independent processes may develop similar solutions to a problem of survival. As an example, the eye of the octopus (a cephalopod) and the eye of man (a mammal) evolved completely independently of

one another. But, as Blakemore (1977) observed, the eyes evolved to be very similar to one another. Both organisms developed functionally similar solutions to the problem of seeing. Blakemore argues that analogously two languages may have developed independently, even though they have properties in common. Following this logic then, the assumption that reading and listening can be viewed as similar processes does not necessitate an assumption of a common phylogenetic or ontogenetic evolution. Given that reading and listening solve the same problem, it is not unreasonable to assume that they are analogous, rather than hierarchical, solutions to the problem.

Analogous solutions to language understanding would not be limited to reading and listening. Consider the recent proliferation of research on the processing of manual-visual languages. On every dimension that has been explored, remarkable parallels have been found between understanding signs and understanding speech. Lane, Boyes-Braem, and Bellugi (1976) found that perceptual confusions among signs can be described utilizing a distinctive feature system, analogous to systems developed for perceptual confusions in speech. Tweney, Heiman, and Hoemann (1977) demonstrated that grammatical structure in sign language plays the same functional role that it does in spoken language. These results support the claim that grammatical processes of language are relatively general and abstract--not tied uniquely to the input modality. The next generation of language-processing theories will have to incorporate the work on processing manual-visual languages. Although this paper is limited to reading and listening, the work on ASL encourages the belief that there are similar and analogous processes in all forms of language understanding.

A STAGE MODEL OF LANGUAGE PROCESSING

Reading can be defined as the abstraction of meaning from printed text; listening as the abstraction of meaning from speech. Deriving meaning from a spoken or written message requires a series of transformations of the energy signal after it arrives at the appropriate receptors. Language processing can be studied as a sequence of processing stages or operations that occur between the energy stimulus and the meaning. In this framework, language processing can be understood only to the extent that each of these processing stages is described. In a previous effort a general information processing model was utilized for a theoretical analysis of speech perception, reading and psycholinguistics (Massaro, 1975a). The model was used heuristically to incorporate data and theory from a wide variety of approaches into the study of language processing. The model should be seen as an organizational structure for the state of the art in language processing. In this paper, I will pre-

sent a general overview of the information-processing model, and use the model to describe and incorporate research that has been carried out in the last few years.

Figure 1 presents a flow chart of the temporal course of reading and listening. At each stage the system contains structural and functional components. The structural component represents the information available at a particular stage of processing. The functional component specifies the procedures and processes that operate on the information held in the corresponding structural component. The model distinguishes four functional components: feature detection, primary recognition, secondary recognition, and rehearsal-recoding. The corresponding structural components represent the information available at each of these stages of processing. The model will be described in more detail below and contrasted with other models of language processing.

The primary assumption of a stage model is that, at some level, the stages are necessarily successive. The feature-detection process transforms the energy pattern created by the language stimulus and transduced by the appropriate receptor system into a set of features held in preperceptual storage (PS). Primary recognition which evaluates and integrates the features in PS into a synthesized percept must necessarily follow feature detection. However, not all of the features must be detected before primary recognition begins. Given that different features require different detection times, not all of the features will enter PS at the same time. Primary recognition could, therefore, evaluate, integrate, and continually update its analysis of the contents of PS while feature detection goes to completion. As can be seen in this example, a basic assumption in the model is that the successive processes necessarily overlap in time. Although this assumption complicates the traditional serial non-overlapping-stages model, the model continues to have predictive and descriptive power. The serial model is necessary, in the present view, because later stages of processing do not have access to early low levels of information at early stages. As an example, rehearsal-recoding does not appear to have direct access to the features that are detected and held in preperceptual storage.

Criticisms of Stage Models

The model under discussion here can be classified in a successive-stage information-processing framework; it therefore has some similarities to the stage models developed by Estes (1975), Gough (1972) and LaBerge and Samuels (1974). Rumelhart (1977) justly criticizes Gough's and LaBerge and Samuels' models because "no higher level can in any way modify or change the

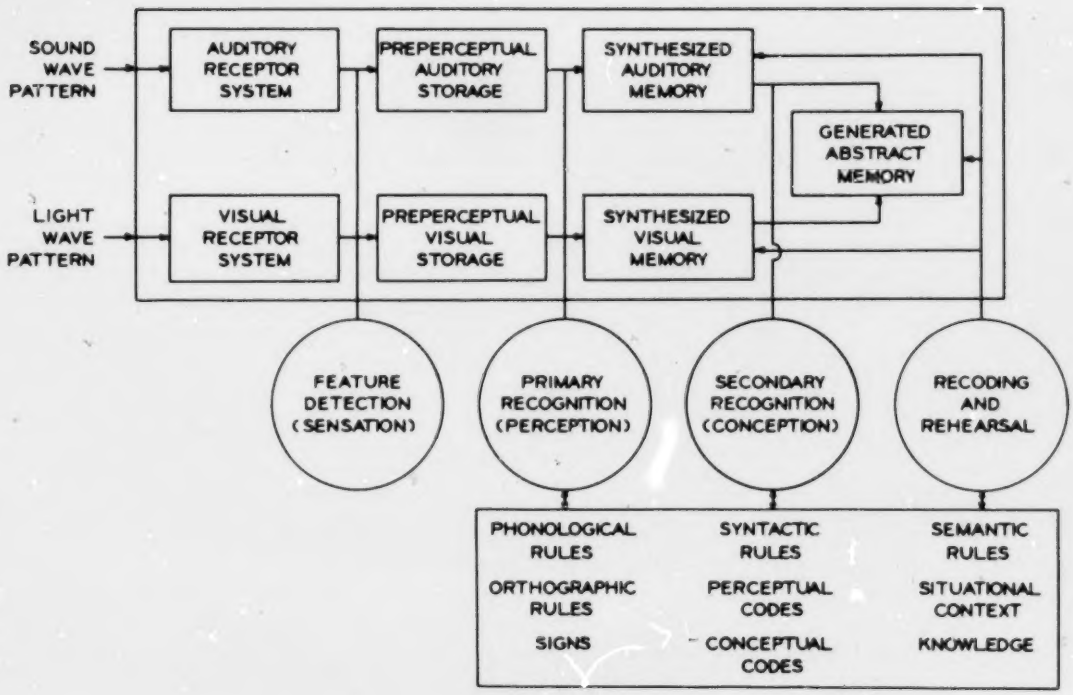

Figure 1. Flow diagram of the temporal course of reading and listening.

analysis at a lower level [p. 578]." As evidence against this assumption, he mentions the well-known finding that the perception of a letter can be influenced by the letters that surround it. A letter that is a part of a word or a pseudo-word is more accurately perceived than the same letter presented in a nonword context. In this case, then, it appears that higher-order knowledge of letter groups or words influences the perception of individual letters. In Rumelhart's opinion and also that of Allport (1977), this result cannot be easily handled by serial models. I would like to stress, however, that not only is this result easily described by the present serial model, the model was also instrumental in generating hypotheses and experiments on exactly this phenomenon (Massaro, 1973; 1975a, Massaro, Venezky, & Taylor, in press; Thompson & Massaro, 1973). In our model two sources of information are available at the primary recognition stage in reading: (1) the featural information in preperceptual visual storage (PVS) and (2) knowledge about the orthographic structure of English spelling stored in long-term memory (LTM). Both of these sources of information contribute to the perception of individual letters in a letter string. (See Massaro, 1975a, and Venezky and Massaro, in press, for a more detailed description of this process.)

Although it is clear that the present model allows orthographic structure to influence letter perception, the influence of syntactic/semantic constraints may not be immediately obvious. It seems obvious that the same visual information can be interpreted differently in different syntactic/semantic contexts (Nash-Weber, 1975). Figure 2 shows that faced with the same visual configuration, the preceding context supports the recognition of cast in one case and east in the other. In our model, knowledge obtained at the level of generated abstract memory (GAM) would be combined with the featural information in PVS. Given the same featural information, different contexts can lead to different outcomes of primary recognition.

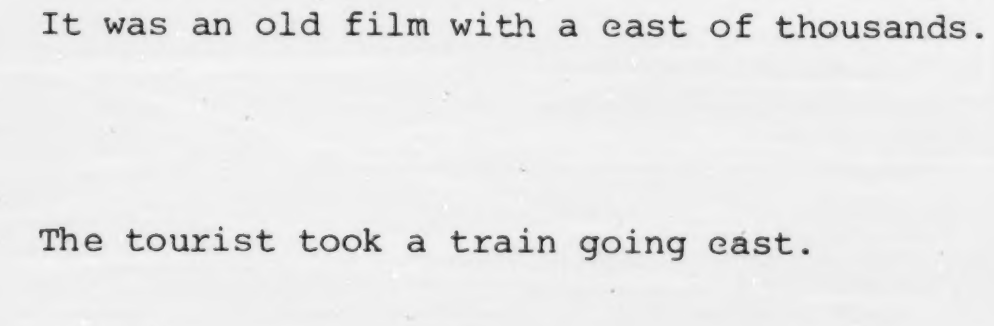

Figure 2. The same visual information can be interpreted differently in different semantic/syntactic contexts.

Although this possibility was acknowledged in our earlier work, very little data were available and, therefore, the influence of semantic/syntactic constraints at primary recognition was not made explicit (however, see Massaro, 1975a, pg. 258-259). More recently, Marslen-Wilson and Welsh (in press) have found that the message being understood at the level of GAM can influence the primary recognition of the message. We will extend the model by developing a quantitative formulation of how syntactic/semantic constraints and featural information are combined and integrated at the primary recognition stage of processing (see the section on secondary recognition).

Parallel Interactive Models

In two recent schemata of reading systems, Allport (1977) and Rumelhart (1977) have drawn heavily and freely from contemporary work in artificial intelligence (AI). The distinctive feature of Rumelhart's system is a set of independent knowledge sources, each of which applies specialized knowledge about some aspect of reading. More specifically, each knowledge source generates hypotheses about the input and relays these to a message center which maintains a running list of all incoming hypotheses. The knowledge sources includes featural, letter, letter-cluster, lexical, syntactic, and semantic levels of analysis. The knowledge sources evaluate their current hypotheses and generate new ones until some decision can be reacned. In this model, all levels of knowledge interact and each contributes independent evidence for the final interpretation.

Although this view has a number of appealing properties, there are also a number of deficiencies when it is evaluated in terms of a psychologically real description of reading. The primary problem is that the model is a direct application of a nonhuman system. In this case, the system HEARSAY II was developed for automatic speech understanding (Lesser, Fannell, Erman, and Reddy, 1975). Although the system proved to be very effective in understanding speech, it had to perform in a manner that was clearly different from the way that humans understand speech. The specific advantage that HEARSAY II has over its human counterpart is that it has a high resolution memory. In HEARSAY II the parametric level holds the most basic representation of the utterance in terms of fine grained acoustic parameters. These low-level data are always available as higher-order hypotheses are made available and tested. In this scheme, low-level decisions are not final since the original parameters can be processed and evaluated continuously throughout the understanding process. By failing to make final low-level decisions, the number of hypotheses and the amount of data that have to be maintained would far exceed the capacity of human memory struc-

tures. We know that preperceptual auditory storage holds featural information for roughly a quarter of a second--less than the time required to hold a complete sentence until it is finally interpreted. This observation is not meant as a criticism of automatic speech recognition, but rather as a caveat for those who seek psychological models in that arena. Although HEARSAY II is a very successful system, HEARSAY II does not make any final decisions about the first part of a sentence until all of the sentence has been processed. Humans recognize speech while it is being spoken if not before. My criterion for the application of AI theories to humans is not that machines perform as well as humans but that machines perform like humans.

Allowing low level data to be continuously available may be more compatible with reading that with speech perception, since the reader can always make a regressive eye movement to refresh the image. In this case, the low level information in reading can be maintained indefinitely as it is in automatic speech recognition. If we consider the number of hypotheses that might have to be checked at this level, however, the number of regressive eye movements would be enormous and our reader would never get through the message. The small number of regressions in normal reading informs the idea that low level information in reading is continually rechecked until a decision is made. Although some concepts developed in AI understanding systems have obvious relevance to psychological theory, current AI schemes are not good models of what we know about how humans process language.

SUMMARY

The model presented here developed in the genre of a serial stage and information processing framework. A number of recent statements have pointed out limitations to this approach (Allport, 1977; Rumelhart, 1977) and have argued for a system based on parallel and interactive channels. In the serial-stage model, each stage has contact with only the information available to that stage and some neighboring stages, and with the appropriate information in long-term memory. For example, the secondary recognition operation in our model has contact with synthesized memory but not with preperceptual storage. Accordingly, the information that was lost during the primary recognition transformation cannot influence secondary recognition. In an interactive model, the featural information at preperceptual storage would influence very late and high-level processing stages. This assumption is unreasonable because the early low-level information doesn't remain intact long enough to be assessed at later stages of processing. For example, an eye movement in reading erases and replaces the featural information derived in

the earlier fixation. It is only the transformed information, then, that can contribute to later processing. The temporal limitations on retention of the various information sources is the primary motivation for utilizing a serial model and in this regard the parallel interactive models are not good psychological models. The stage model will now be developed in more detail by tracing the flow of information processed in reading and listening. By reviewing relevant issues and research, it is hoped that the heuristic value of the model will become apparent.

II

FEATURE DETECTION

The feature-detection process transforms the energy pattern into features in preperceptual storage. The features for reading and listening are described as visual and acoustic, respectively, since we assume that there is a direct relationship between the nature of the energy signal and the information in preperceptual storage. There are a number of prominent characteristics in speech and reading stimuli that are potential candidates for features. A characteristic is called a feature only when it is psychologically functional, that is, when it is used to distinguish between different speech sounds or written symbols.

AUDIBLE FEATURES

The study of features in speech perception has a long history. A good portion of the work was influenced by linguistic descriptions of binary all-or-none distinctive features (Jakobson, Fant, & Halle, 1961). One of the goals of distinctive feature theory was to minimize the number of distinctive features of the language. Distinctive features were designed to be general and not stimulus specific: that is, if a distinctive-feature difference distinguished two phonemes in the language, that same distinction was assumed to distinguish several other phoneme pairs. Given the distinctive feature of voicing, for example, the distinction of voiced versus voiceless could account for the differences between /z/ and /s/, /v/ and /f/, and so on.

In contrast to the linguistic description of binary all-or-none features, the features held in preperceptual auditory storage (PAS) are assumed to be continuous, so that a feature indicates the degree to which the quality is present in the speech sound. Rather than assuming that a feature is present or absent, it is described in terms of its degree of presence in PAS. This assumption is similar to the more recent treatment of distinctive features provided by Chomsky and Halle (1968) and Ladefoged (1975). Chomsky and Halle distinguish between the classificatory and phonetic function of distinctive features. Features are assumed to be binary in their classificatory function, but not in their phonetic or descriptive function. In the latter, features are multivalued representations that describe aspects of the speech sounds in the perceptual representation. Ladefoged distinguishes

between the phonetic and phonemic level of feature description. A feature describing the phonetic quality of a sound has a value along a continuous scale whereas a feature classifying the phonemic composition is given a discrete value. The assumption of continuous representations of acoustic features is, therefore, exactly analogous to the linguistic representations of phonetic quality.

In the present model an acoustic feature in preperceptual auditory storage (PAS) is expressed as a continuous value. It follows that the listener will be able to hear the degree of presence or absence of a particular feature, even though his judgment in a forced choice task will be discrete. Oden and Massaro (1977) describe acoustic features as fuzzy; that is, varying continuously from one speech sound to another. In this representation features are represented as fuzzy predicates which may be more or less true rather than only absolutely true or false (Zadeh, 1971). In terms of the model, fuzzy predicates represent the feature detection and evaluation process; each predicate is applied to the speech sound and specifies the degree to which it is true that the sound has a relevant acoustic feature. Rather than assuming that a sound is voiced or voiceless, for example, the voicing feature of a sound is expressed as a fuzzy predicate

$$P(\text{voiced}(S)) = .75 \qquad (1)$$

The predicate given by Equation 1 states that it is .75 true that speech sound S is perceived to be voiced. In terms of the model, then, the feature detection process makes available a set of fuzzy predicates at the level of PAS. In addition to being concerned with the acoustic features in preperceptual storage this analysis of the feature-evaluation process makes it apparent that an important question in speech perception research is how the various continuous features are integrated into a synthesized percept.

The integration of features has not been extensively studied in speech perception research for two main reasons. If distinctive features are binary, the integration of information from two or more features would be a trivial problem. Given binary features representing voicing and place of articulation, for example, the integration would be a simple logical conjunction. If the consonant /b/ is represented as voiced and labial, a voiced labial sound would be identified as /b/ whereas a voiceless labial sound or a voiced nonlabial sound would not be identified as /b/. A second reason for the neglect of the integration problem is methodological. The primary method of study involved experiments in which the speech sound was varied along a single relevant dimension. For example, in a study of voicing all voicing cues were made neutral except one, such as voice onset time and then this dimension was varied through the relevant values. Similarly,

in the study of place of articulation all cues but one were neutralized, and the remaining dimension was varied through the appropriate values. Very few experiments independently varied both voicing cues and place cues within a particular experiment. Therefore, no information was available about how these cues were integrated into a synthesized percept. More recently, Massaro and Cohen (1976) and Oden and Massaro (1977) have utilized factorial designs and functional measurement techniques (Anderson, 1974) to study the integration of acoustic features in speech perception. The work has illuminated how fuzzy features are integrated into a synthesized percept. The methods allow the investigator to determine which acoustic characteristics in the signal are perceptually functional, the relative weight that each feature carries in the integration process, and exact combinatorial algorithm that is used.

VISIBLE FEATURES

One of the oldest areas of reading-related research is the study of the functional cues used in recognizing printed characters. Much of this work was performed by typographers and artists concerned with the relative merits of good design and good legibility in type fonts. Although many of the early conclusions remain valid today, they are almost totally ignored in the contemporary study of letter recognition. The primary influence in extant studies has been the well-known neurophysiological findings that the responses of cells in the visual cortex demonstrate an amazing stimulus selectivity (Hubel & Wiesel, 1962). For example, there appear to be specialized detectors in the visual system for lines of specific size and orientation (see Blakemore, 1973 and Lindsay & Norman, 1977 for intelligible reviews of this work).

Consistent with the all-or-none response of the specialized cells, psychological descriptions have centered around binary all-or-none features. Feature sets usually consist of the presence or absence of horizontal, vertical, or oblique lines, curves, intersections, angles, and so on. The feature sets are typically derived from and tested against the recognition confusions of capital letters (see Massaro, 1975a, Chapter 6). In contrast to the idea of binary all-or-none features, however, visible features, like audible features, may be fuzzy. Rather than a feature being present or absent, the information in preperceptual visual storage (PVS) could represent the degree to which a given feature is present in the signal. Given this conceptualization, it is necessary to develop a methodological and theoretical study of visible features that parallels the work on audible features (see the section above on Audible Features).

Given the idea of fuzzy information, it is important to carry out research that manipulates the degree to which a feature

is present in a letter. Recently, Blesser and his colleagues have developed and studied ambiguous characters (Blesser, Shillman, Kuklinski, Cox, Eden, & Ventura, 1974; Shillman, Cox, Kuklinski, Ventura, Blesser, & Eden, 1974). Figure 3 presents a matrix of ambiguous characters having one or more functional attributes in

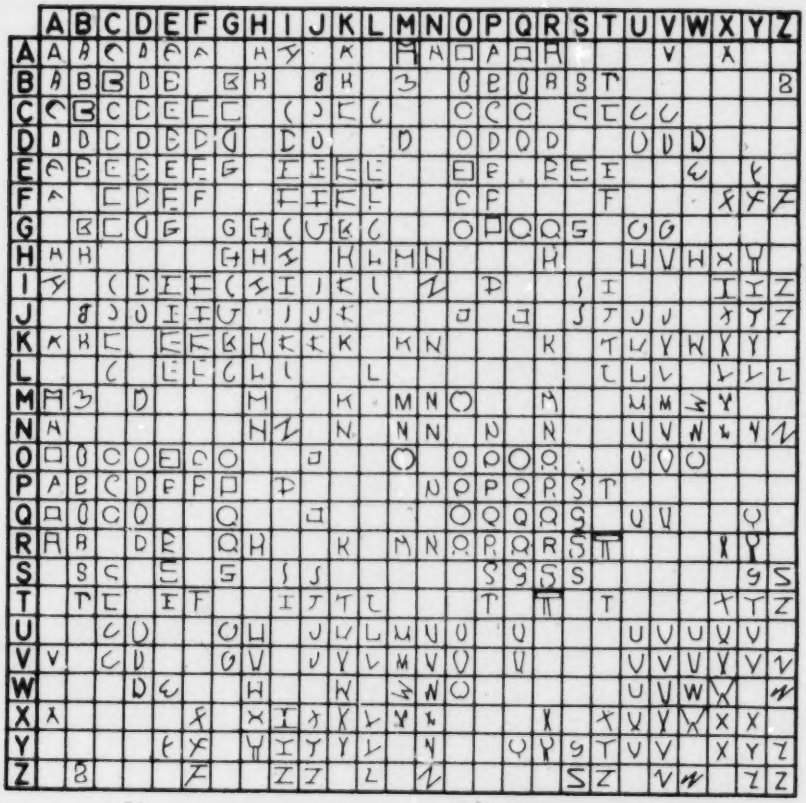

Figure 3. A matrix of ambiguous characters having one or more functional attributes in transition between letter pairs (after Naus & Shillman, 1976).

transition between two letters. A completely ambiguous character is one that would be assigned to either letter class with equal probability. As an example, a v can be gradually transformed into a y by continuously increasing the right oblique line below the intersection (Naus & Shillman, 1976). This work with ambiguous characters is consistent with the idea of fuzzy visible features, since each feature must be represented in terms of the degree to which it is present in the character. Analogous to the recent work in speech perception, the theoretical notion of fuzzy information and the experimental methods of factorial designs and functional measurement techniques should advance the study of visible features in reading.

Given that most of the contemporary feature perception research is carried out with uppercase letters, it is important to know to what extent these features may generalize to the perception of lowercase letters. Bouma (1971) showed that the visual confusions in the recognition of lowercase letters were better described in terms of the overall shape or envelope of the letter than by the commonly-assumed component features. The envelope of a letter is defined as the smallest enclosing polygon without indentations. The envelope for e would be circular, v would be enclosed by an inverted triangle, and so on.

Bouma's results indicate that the overall envelope of a letter may be resolved before details of its component features. Independently of whether the whole is greater than the sum of its parts, a general impression of the whole (although ambiguous) is seen before all of the parts. That overall envelope of the letter becomes available before the letter details is also consistent with fourier-analysis studies. There is good evidence that information given by low spatial frequency is processed faster than high frequencies (Breitmeyer & Ganz, 1976; Broadbent, 1977). Accordingly, low-frequency information could specify envelope shape before high-frequency information reduces uncertainty about letter detail. Either limited processing time or a reduced figure-ground contrast could allow envelope resolution without corresponding resolution of letter detail. In terms of the present analysis, the reader will have to decide which letter best fits the perceived but fuzzy envelope. For example, recognition of a circular envelope without further detail would be sufficient to limit the alternatives to the three letters e, c, and o.

VISIBLE FEATURES IN WORDS

There has been an abundance of research work on letter legibility, which is nicely summarized by Spencer (1968). Javal (cited in Spencer) showed that reading is much easier when the top half, rather than the bottom half of a line of words is exposed. It is not clear where the vertical center was defined but we might assume that it was in the middle of an x-height

letter. It would be valuable to define the perceptual half of words; i.e., the point at which reading the top and bottom portions of a line of print is equally easy. Consistent with Javal's finding, Messmer's research led him to conclude that ascending letters are dominant and contribute more to word recognition than x-height or descending letters. Zeitler also found that ascenders were dominant. In short word presentations, letters projecting above or below the line were more likely tc be perceived (see Huey, 1908, pp. 79-87). It would be interesting to repeat Javal's experiment with a type font that equates the vertical extent for all letters, so that tails do not extend above or below the line. This font was actually developed by Andrew Tuer in the 1880's and is only now being used in some computer printouts of lowercase type.

III

PRIMARY RECOGNITION

In the framework of the information processing model, the primary-recognition process integrates the featural integration held in preperceptual storage into a percept in synthesized memory. One question relevant to the model concerns the functional units at this stage of processing. The functional units are called perceptual units; they correspond to units that are described in long-term memory. The primary-recognition process finds the best match between the featural information and the descriptions of perceptual units in long-term memory.

PERCEPTUAL UNITS IN SPEECH

Sound patterns of V, CV, or VC size are the best candidates for perceptual units since these patterns can be described by relatively invariant acoustic features (Fujimura, 1975). Smaller patterns such as phonemes lack invariance, and in fact, this lack of invariance has been one of the central focuses of speech perception theory (see Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). Although perceptual units of larger size have also been proposed, research has demonstrated that preperceptual auditory storage cannot maintain featural information for a period of time greater than roughly one quarter of a second. It follows that clauses, phrases, or even words are inappropriate perceptual units at the primary recognition stage of processing. (For further discussion see Massaro, 1975a, Chapter 4.)

Target Search in Listening

If syllables are perceptual units in speech processing, they should be accessed before the phonemes that make them up. One test of this idea has been to study the reaction times to syllable and phoneme targets in speech processing. Savin and Bever (1970) asked listeners to release a telegraph key as soon as they heard a specified target in a sequence of nonsense consonant-vowel-consonant syllables (CVCs). Listeners were slower in responding to initial consonant phoneme targets such as /b/ or /s/ than in responding to a CVC target such as /bæb/ or /sæb/. They interpreted the differences in reaction times (RTs) as directly reflecting differences in the time to perceive the phoneme and

syllable targets, and concluded that phonemes must be perceived only by an analysis of already perceived syllables. Warren (1971) researched similar conclusions from a target-search task through spoken sentences.

Foss and Swinney (1973) objected to the conclusion that phoneme perception must follow syllable perception. They used search lists of two-syllable words and asked subjects to search for an initial phoneme, initial syllable, or a complete word. Reaction times were actually somewhat shorter for word than for syllable targets, although phoneme targets were responded to slowest of all. In terms of the logic of the target-search task, their results would argue that words are accessed before the syllables that make them up. Burying the logic of the target-search task even deeper, McNeill and Lindig(1973) included sentence targets and lists in a complete factorial design of phoneme, syllable, word, and sentence targets crossed with the same four kinds of lists. In general, RTs were slowed when the linguistic level of the target differed from that of the search list.

The latter studies have been used to support the thesis that the target-search task is inappropriate for a study of perceptual units in speech processing. This conclusion may be premature, however, given the inappropriate design of the experiments. The critical confounding was that appropriate test foils were not included in the experiments. For example, Foss and Swinney's (1973) results indicated that subjects were not faster in responding to syllable versus word targets. However, the foils were chosen in such a way that subjects could have based their decision on the first syllable of the word in both the target syllable and target word conditions. Given either the syllable target can or the word target candy, subjects could have selected a reliable response after processing just the first syllable, since the test foils never began with can. Apparently this is the strategy that was used since the RT averaged about 350 msec, much less than the 500 msec duration of the test words. Similarly, McNeill and Lindig's subjects responded in less than 1/3 sec, showing that a response was initiated long before a word or sentence could have been processed. Therefore, the target search task should be reconsidered as an appropriate paradigm, but it should be structured to account for both perceptual and comparison processes (see Ball, Wood, & Smith, 1975; Massaro & Kiltzke, 1977; Sloboda, 1977).

PERCEPTUAL UNITS IN READING

In contrast to the acoustic structure of phonemes, printed letters of a given font are invariantly defined with respect to their visual characteristics. It follows that units as small as letters are reasonable candidates for perceptual units at the

primary recognition stage in reading. Units of larger size have been proposed, however. Spelling patterns, pronunciation units, words, and even phrases have been proposed as perceptual units at the primary recognition stage of processing(the integration of visual features into a synthesized percept).

If letters are basic perceptual units in primary recognition, words must be perceived by way of the letters that make them up. This hypothesis addresses the old and familiar question of whether word recognition can be described in terms of component letter recognition or whether a word is recognized on the basis of supra-letter features without reference to the letters that make it up. This question has generated more than enough interest in the last decade, but some recent experimental findings that have been used to reject the hypothesis that letter perception mediates word perception are of interest in this discussion.

If words are recognized via the letters that make them up, Cosky (1976) reasoned that the ease of word recognition should be a direct function of the ease of recognition of the component letters. He measured the time to name letters presented alone and the time to discriminate between letters and used these times as an index of letter legibility. Then he composed words of the 13 most legible letters and the 13 least legible letters. Employing these words in a naming task, he found no effect of letter legibility and concluded that letter perception does not mediate word perception. However, in addition to interpreting differences in naming times as a direct index of differences in recognition time (see Chapter IV, Phonological Mediation) there are a number of limitations to Cosky's study. Most importantly, Cosky (1976) did not show that his grouping of letters in terms of their legibility could predict performance in multi-letter items that are not words. If the uniqueness of words is responsible for Cosky's negative finding, then positive results should occur in nonword strings. Until this is demonstrated, Cosky's results can only be taken as a failure to find proof for the letter-unit model; it cannot be taken as disproof.

Target Search in Reading

The target-search task has also been used in reading to address the issue of whether word recognition is mediated by letter perception. In an experiment carried out by Johnson (1975), subjects indicated whether or not a test item was the same as a target item. Response latencies did not depend on whether the target and test items were single letters or whole words. These results led Johnson to reject the idea that word recognition is mediated by letter recognition. However, Johnson's results are not necessarily incompatible with preliminary letter recognition if the letters in a word can be processed in parallel.

Massaro and Klitzke (1977) observed that the population of words of a given length may be more dissimilar from one another than are a set of single letters. If this is the case, a subject could perform with a less stringent criterion for accepting a match in the word condition that in the letter condition. Johnson (1975) made no attempt to control for the similarity of the test foils to the target items thereby producing test foils that were naturally more dissimilar on word than on letter trials. Increasing the similarity between the target and test foils should decrease the advantages of parallel processing on word trials and, therefore, produce a letter advantage because of reduced figure-ground contrast (lateral masking) and limited capacity effects in processing word strings.

Massaro and Klitzke (1977) replicated Johnson's experiment while simultaneously investigating the role of similarity in the identification of words and letters. Half of the nontargets were highly similar to the target and half were dissimilar. The manipulation of similarity was effective in that responses to similar nontargets required 40 msec more than responses to dissimilar nontargets. In contrast to Johnson's finding of no difference, reaction times were 8 percent (43 msec) faster for letter stimuli than word stimuli. The letter advantage was significant even when the analysis was restricted to target and dissimilar nontarget trials. By including similar nontarget trials, subjects could not maintain a lax criterion of sameness because a nontarget test word could be very similar to the target word. In this case, subjects could not select and execute a positive response as quickly as they might have with a more lax criterion of sameness. Massaro and Klitzke interpret the results as suggesting that perceptual processing and comparison are continuous and overlapping processes in the target-search task. Subjects do not first perceive the test item and then compare it to the target item in memory but rather the subjects are able to make comparisons as partial information about the test item is resolved. Notice that this interpretation is consistent with the general idea of successive but overlapping stages in language processing. Therefore, the normal dissimilarity advantage of word strings over single letters can work to neutralize the processing advantage of single letters. Another advantage for words is orthographic structure (see below, letter vs. word units). But in both cases the word advantage is not incompatible with idea of letter units mediating word recognition (see also Sloboda, 1976, 1977).

The Word Advantage

In a well-known experiment carried out by Reicher (1969), subjects were presented with either a single letter, a four-letter word, or a four-letter nonword flashed in a tachistoscope

and had to report what they saw. Reicher's contribution to this century-old task was to constrain the subject's choice by presenting two one-letter alternatives after each trial. Both alternatives completed the display by spelling words in the word condition so that performance on the word trials would not benefit from a simple guessing strategy. Even with these constraints, Reicher found a 10 percent advantage for recognition of a letter in a word, over recognition of a letter in a nonword or a letter presented alone. These results have been described both in terms of single-letter and whole-word perceptual units. In terms of a perceptual unit the size of a word, the word has supraletter features such as overall word shape which facilitate direct contact with the appropriate memory representation. A word is recognized better than a letter or a nonword because the unique visual features of a word allow for easier recognition than the features of a single letter or a nonword.

The advantage of words over single letters and nonwords is not incompatible with the idea that the letter is a basic perceptual unit, however. In this model, the primary recognition process operates on a number of letters in parallel. The visual features read out at each letter position define a candidate set of possible letters for that position. The recognition process is not limited to featural information, but can also utilize knowledge about the orthographic structure of English spelling. The letter than is synthesized at each position, therefore, will not only correspond to the visual information that is available from feature detection and evaluation, but will also correspond to the orthographic constraints in the language. For example, consider the case in which the subject is given the lowercase string coig and has resolved just the circular envelope of the first letter and all of the last three letters. Given that c, e, and o are the only letters that are consistent with the circular envelope, these are the only possible letters at this position. If the reader further assumes that the string must conform to English orthography, only c is possible since the strings ooig and eiog are illegal English spellings. In this case, the reader can synthesize coig since it is the only valid alternative. When the single letter c is presented, on the other hand, perception of the envelope does not allow an unambiguous choice between c, e, and o. Accordingly, the reader is equally likely to synthesize any of the three letters and will be correct only one out of three times. Although a word is recognized via its component letters, familiarity with the orthographic structure of words facilitates primary recognition of the letters relative to a single letter or nonword.

Single Letters vs. Word Units (An Experimental Test)

One way to distinguish between these two interpretations of the word advantage is to demonstrate that letter perception is important

in word perception. If the letter-unit idea is correct, factors that determine letter perception in single letters and nonwords should also determine letter perception in words. To the extent that a letter-unit model can describe the recognition of words, nonwords, and letters in terms of the same processes, a word-unit model is not warranted by the data. Given this outcome, the letter unit model should be preferred until an equally parsimonious word-unit model is articulated. Massaro and Klitzke (unpublished) showed that two relevant factors in single letter and nonword perception operate identically in word perception. They are (1) the processing time available between the onset of a test letter and the onset of a masking stimulus and (2) the lateral masking of the perceptibility of a letter by its neighboring letters.

A quantitative formulation of the primary recognition stage describes the temporal course of letter perception by the simple equation.

$$d' = \alpha(1-e^{-\theta t}) \qquad (2)$$

where d' represents the perceptibility of the letter in z-units, α represents maximal perceptibility with unlimited processing time, t the processing time measured from the onset of the test stimulus to the onset of a masking stimulus, θ the rate of processing, and e the natural logarithm. Perceptibility is assumed to be a negatively accelerating growth function of the processing time available. The value α is dependent on the properties of the visual display and the acuity of the visual system. The value θ reflects the rate of perceptual processing--how quickly primary recognition occurs. The value of θ can be expected to be dependent upon process variables such as selective attention and the degree to which the reader utilizes orthographic regularity.

The primary support for this description of the temporal course of letter recognition comes from backward recognition masking experiments. In this task, a target stimulus is presented for a short duration and is followed after some blank interval by a masking stimulus superimposed at the same location as the target. The basic finding is that recognition of the target improves with increases in the blank interval before the onset of the mask. Before applying the model to the issue of letter and word units, some clarification of backward masking is necessary. There seems to be a general misunderstanding of what backward masking does and how it is interpreted in serial-stage models of information processing. One common misinterpretation is that a backward mask somehow works retroactively: it eliminates all perceptual information about the visual display. Unless the subject managed to encode the display into an abstract representation, he will not be able to report

anything about what was presented. In our view of backward masking, however, the mask simply terminates any further perceptual resolution; the resolution that occurs before the mask is presented is continuously passed on to the next stage. This information is not eliminated by the mask, although the mask may also function to interfere with the information at this level (see Massaro, 1975b, Chapter 24). The gradual improvement in performance with increases in the masking interval reflects the continuous perceptual resolution before the mask is presented. For a further discussion of backward masking, see Massaro (1975b, Chapters 17, 18, 22, and 23).

In the model, lateral masking and orthographic regularity are identified with different parameters in the equation describing performance as a function of processing time t. Adjacent letters that degrade the perceptibility of a neighboring letter should influence the α value for that letter. Adjacent letters that reduce the uncertainty of a given letter should increase θ, the rate of processing the information in that letter. Consider the case in which the single letter c is presented. Perceptual resolution of the letter is assumed to be a temporally extended and continuous process. As an example, when t = 100 msec, the letter may be resolved sufficiently to reduce the alternatives to c, e, and o; 200 msec might be needed to completely resolve the letter c. In the analogous word case, the letter c may be presented in coin. In this case, because of lateral masking c may not be resolved completely but may be seen at only 90 percent clarity even with unlimited process time. Hence, the α value for c would be lower in the word than in the single letter condition. But the word context should also enhance the rate of processing the information in the letter c. If the context oin is resolved completely and the alternatives for the first letter are limited to c, e, and o, no further visual processing is necessary given this visual information and the orthographic regularity of the written language. The strings eoin and ooin are illegal in the context oin. Therefore, c is the only valid alternative for the first letter. Given limited visual information, c can be recognized exactly in the word context but not when it is presented alone.

Using the rules of orthographic regularity, the subject can arrive at a correct decision more quickly in a word than in a single-letter condition. The operation of these two processes predicts an interaction between the letter and word conditions and the processing time available. When processing time is maximal, rate of processing is unimportant and the letter condition should show an advantage because the perceptibility of a letter in a word is reduced by lateral masking. With intermediate processing times, the advantage of orthographic regularity in the word condition should enhance the rate of processing of the test letter and, therefore, offset the deficit of lateral masking. In the Massaro and Klitzke experiments processing time was

controlled in a backward recognition masking task in which the test display was followed by a masking stimulus after a variable blank interval. Single letter, words, nonwords, and letters embedded in dollar signs were used as test items in the Reicher (1969) task. The two test alternatives, always presented 250 msec after the test item, were the complete test item and its corresponding foil that differed by one letter. Performance on nonwords should be equivalent to that on letters in dollar signs at all masking intervals. Both have the disadvantage of lateral masking and both possess no orthographic regularity. The model also predicts that the disadvantage of lateral masking in words relative to single letters (different α's) can be overridden at the shorter mask onset intervals by the faster rate of processing letters in words (different θ's) because of the utilization of orthographic regularity.

The result of critical interest to the letter versus word unit models is the interaction of context and processing time. In order to test the model, d' values were computed for the letter, word, and nonword conditions as a function of stimulus onset asynchrony, the time between the onset of the test display and the onset of the masking stimulus. The d' values were computed from the average percentage correct values. In this analysis, the hit rates correspond to the percentage correct values and one minus these values are the false alarm rates. The nonword and the letter-in-$ conditions were also averaged before the d' values were computed, since there was no significant difference between these conditions.

Figure 4 plots the d' values as a function of context and the stimulus onset asynchrony. Although a monotonic masking function was observed for each context condition, the functions appear to rise at different rates and to different asymptotes in the different conditions. Although the perceptibility of a letter in a word is greater than a letter presented alone at short processing times, the opposite is the case for long processing times. There is no easy way that a whole-word perceptual unit can explain this result. In contrast, this interaction is exactly what is predicted by the present formulation of the letter-unit model. In terms of the model, α should be larger for single letters than words or nonwords because of the reduced signal-to-noise ratio due to the lateral masking of adjacent letters in the word and nonword conditions. The rate parameter θ, however, should be larger for words than for single letters or nonwords since orthographic constraints allow the reader to arrive at a decision about which letters are present at a faster rate in the word condition than in the single-letter condition. The clarity of the features for a letter in a word is less than for a letter presented alone but fewer of these features are necessary to arrive at a decision in the word condition than in the single-letter condition.

In fitting the model to the results, one α was estimated for

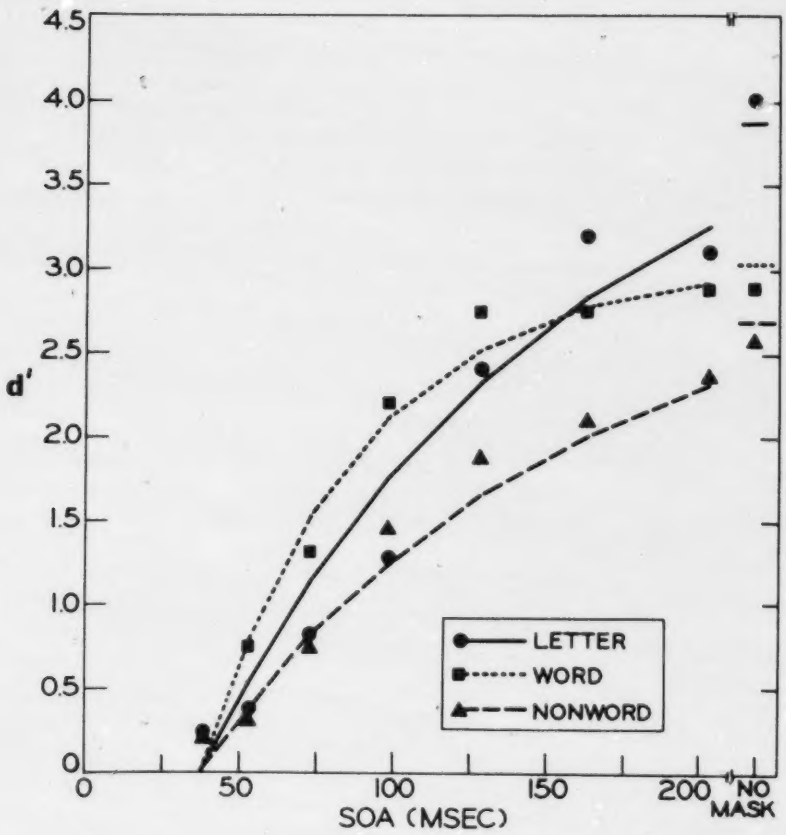

Figure 4. Predicted and observed d' values for the word, nonword, and single letter conditions as a function of SOA.

the single-letter condition and another for the word and nonword conditions. The word and nonwords should have the same value of α since lateral masking should be equivalent in these two cases. With respect to the rate of processing θ, letters in words should be processed at a faster rate than letters in nonwords or a letter presented alone. Therefore, one value of θ should be estimated for words and another for letters and nonwords. It was also necessary to estimate a dead time since the masking stimulus was more intense than the test stimulus. Therefore, the masking interval probably overestimated the true processing interval since the mask would have a faster arrival time at PVS than would the test stimulus. Finally, it was necessary to estimate the duration that the display information was maintained in preperceptual visual storage (PVS), since it could not be available indefinitely. This duration, t_D, gives the available processing time when stimulus onset asynchrony is longer than t_D. If the masking interval exceeded t_D, then t_D was inserted in the equation.

The observed d' values were fit with the predictions of the model by estimating the six parameter values using the minimization subroutine STEPIT (Chandler, 1969). The α value for the letter alone condition was 4.32--significantly larger than the α value of 3.06 for words and nonwords. The θ value for words was 18.95--significantly larger than the θ value of 8.45 for letters and nonwords. The dead time was estimated to be 37 msec and 266 msec was the estimated duration of PVS. The model provided a reasonably good description of the results, considering that 24 independent data points were described with just 6 parameter values. The average squared deviation between the predicted and observed values was .038. This fit was not significantly improved when the constraints on α and θ were removed and two additional parameters were estimated. The latter description gave an average squared deviation of .034.

One possible limitation of the present interpretation is that the differences in the context conditions were observed at different levels of overall accuracy in the task. The word advantage at short stimulus onset asynchronies (SOAs) was observed when performance was poor whereas the letter advantage at long SOAs occurred with good overall performance. To eliminate the possibility that these results are unique to performance level, the first experiment was replicated but average performance was maintained at a fixed level at each masking condition. If the interactions between context and processing time are not due to the different levels of overall performance, the same results should be observed when average performance is constant at each masking condition. And in fact the results were replicated exactly: a word advantage was observed at short-processing intervals whereas a letter advantage was observed at long-processing intervals. These experiments offer nice support for a letter-unit model. It is difficult to envision how a whole-word

unit model can be modified to account for the results.

A Second Critical Test

Spelling patterns, pronunciation units, vocalic center groups, words, and even phrases have been proposed as perceptual units at the primary recognition stage of processing. On the basis of parsimony the smallest perceptual unit is preferred, but the answer to the empirical question must come from the extent to which higher or lower perceptual units are functional in reading. The well-known finding is that a letter embedded in an orthographically legal context appears to be recognized better than in an illegal orthographic context. This area of research has skyrocketed to say the least, but all of the results can be interpreted in terms of a model which assumes that letters have perceptual units and that higher order information reduces uncertainty about letter information. The reader is referred to Massaro (1975), Massaro and Klitzke (1977), and Massaro, Venezky & Taylor (in press) for recent work.

One critical question in this area of research is to what extent low-level feature analysis and higher-order context are independent contributions. Krueger (1975) concludes from an exhaustive and careful review of the literature that orthographic structure (familiarity in his view) influences the feature extraction stage. This means that feature resolution should be better in familiar than in unfamiliar letter strings. Other higher order unit models assume that higher-order units intervene in the processing sequence to change in the perceptual analyses which are employed. In Johnson's (1977) model, for example, the feature set assigned to an input is determined by the entire input and higher-order memorial feature sets. In this case, a letter will be assigned different features in different contexts.

In top-down models, higher-order information actually changes the nature of the perceptual analyses. The central assumption of hypothesis-testing models is that the current hypothesis directs perceptual analyses. Perceptual processing and experience is directly dependent on the guiding hypothesis; for example, an hypothesis that the upcoming word is Philadelphia might guide the perceptual analyses to test for an initial capital letter and a relatively long word. Given only these analyses, no other information would be available about letter identity. When hypotheses do not guide perceptual analyses, however, the same letter information would be available regardless of higher-order constraints. A top-down hypothesis-testing model predicts that higher-order information can guide perceptual analyses, whereas an independence model does not. In the latter, the feature evaluation process places the same information in PVS, independently of the higher-order constraints that are available.

Higher-order information operates the level of primary recognition, which synthesizes a percept and passes it on to SVM. Primary recognition can integrate and synthesize the featural information available more quickly and reliably to the extent that higher-order context is available.

An experiment was carried out to test between these two general classes of theories. Higher-order unit models that assume that orthographic context influences feature extraction assume that context and featural processing are not independent. It follows that a model that assumes independent contributions of each must fail. Given that no quantitative formulations of the nonindependence hypothesis are available, we are limited to testing the goodness of fit of an independence model. If it describes the results reasonably well, it offers some evidence against the class of nonindependence theories until an equally parsimonious model can be developed and tested favorably against the results.

The experiment involves the independent variation of visual information and context in a letter perception task. Consider the lower case letters c and e. It is possible to gradually transform the c into an e by extending the horizontal bar. To the extent the bar is long, there is good visual information for an e and poor visual information for a c. Now consider the letter presented as the first letter in the context -oin and the context -dit. Only c is orthographically legal in the first context since three consecutive vowels violate a rule of English orthography. Only e is valid in the second context since the cluster cd is an invalid English pattern. In this case, the context -oin favors c whereas the context -dit favors e. The context -tsa and -ast can be considered to favor neither e nor c. The first remains an illegal context whether e or c is present and the second is orthographically legal for both e and c.

The experiment factorially combined 6 levels of visual information with these 4 levels of orthographic context, giving a total of 24 experimental conditions. The bar length of the letter took on 6 values going from a prototypical c to a prototypical e. The test letter was presented at each of the 4 letter positions in each of the four contexts as shown in Table 1. The test string was presented for a short duration followed after some short interval by a masking stimulus composed of random letter features. The subject was asked to indicate whether an e or c was present in the test display. The dependent variable is, therefore, the proportion of e responses at each experimental condition.

The independence model provides a straightforward interpretation of the situation. Two independent sources of information are available: the visual information from the critical letter and the orthographic context. These sources of information can be indicated by V_i and C_i, respectively. The value V_i specifies how much e-ness is given by the critical letter. This value lies

TABLE 1

CONTEXTS FOR THE c-e TEST LETTER AS A FUNCTION OF THE ORTHOGRAPHIC CONTEXT AND SERIAL POSITION

	Serial Position			
	First	Second	Third	Fourth
neutral	_$$$	$-$$	$$_$	$$$_
e and c illegal	_tsa	a_sr	dt_u	tla_
e and c legal	_ast	s_ar	du_t	tal_
c legal; e illegal	_oin	s_um	pa_k	zin_
e legal; c illegal	_dit	s_ll	sl_d	pan_

between zero and one and is expected to increase as the length of the bar is increased. The value C_i represents how much the context supports the letter e; it too lies between zero and one and should be large when e is legal, small when e is illegal, and intermediate when neither or both e and c are legal. With just two alternatives, it is reasonable to assume that the amount of c-ness is simply one minus the amount of e-ness (Massaro & Cohen, 1976). Therefore, if V_i specifies the amount of e-ness of the test letter, then $(1-V_i)$ specifies the amount of c-ness of that same test letter. Similarly, if C_i represents the amount of e-ness given by the orthographic context, then $(1-C_i)$ specifies the amount of c-ness given by that same orthographic context.

The reader has two sources of information, V_i and C_i. At the evaluation stage, the reader evaluates the amount of e-ness and the amount of c-ness from these sources. The amount of e-ness and c-ness for a given test display can be represented by the conjunction of the two independent sources of information:

$$\text{e-ness} = (V_i \wedge C_i) \tag{1}$$

$$\text{c-ness} = [(1-V_i) \wedge (1-C_i)] \tag{2}$$

It is now necessary to define conjunction in the model so that the e-ness and c-ness values given by both sources can be determined. Previous tests in other situations have shown that a multiplicative combination provides a much better description than an additive combination. If the multiplicative combination is also applied here, Equations 1 and 2 can be represented as Equations 3 and 4, respectively:

$$\text{e-ness} = V_i \times C_i \tag{3}$$

$$\text{c-ness} = (1-V_i) \times (1-C_i) \tag{4}$$

Finally, the independence model must specify how a response is determined from the e-ness and c-ness values. Again, following work in other domains, a choice of e is assumed to be made by evaluating the degree of e-ness relative to the sum of e-ness and c-ness values. This choice rule is a direct application of Luce's (1963) choice axiom. In this case the probability of an e response, P(e) can be expressed as:

$$P(e) = \frac{V_i C_i}{V_i C_i + (1-V_i)(1-C_i)} \tag{5}$$

The important feature of the independence model is that the visual contribution specified by V_i does not change with changes in the orthographic context. The nonindependence views expressed by Krueger, Johnson, and other top-down models reject this

assumption because in essence the whole is greater than the sum of its parts. Applying the model to the results, it is necessary to estimate 6 values of V_i for each level of bar length of the critical letter and 4 values of C_i for the 4 experimental contexts. These values were estimated using the iterative routine STEPIT by minimizing the deviations between the observed and predicted values.

The observed and predicted values are given in Figure 5. The independence model provides a good description of the results. The probability of an e response increased with increases in the bar length of the critical letter and was significantly higher for the context in which e but not c was orthographically legal than the context in which c but not e was legal. The neutral contexts were intermediate and there was no significant difference between the both-legal and both-illegal contexts. The parameter estimates are psychologically meaningful. The value of V_i increased in the length of the critical letter. The values were .14, .14, .24, .66, .88

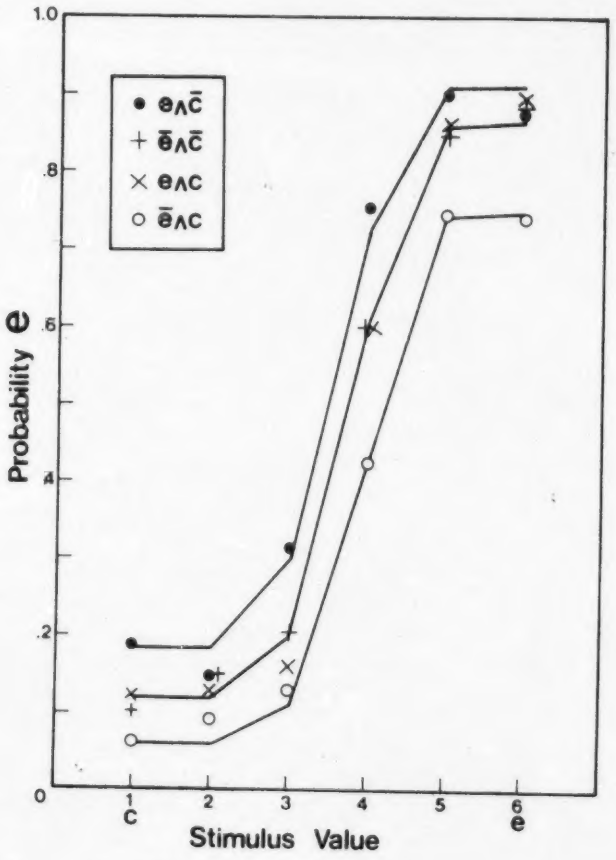

Figure 5. Predicted and observed probabilities of an e response as a function of orthographic context and stimulus value.

and .89 for the six respective levels. The value of C_i was .58 for e legal and c illegal and .28 for c legal and e illegal.

The value of C_i was .45 for the case when both letters were illegal and .45 for the case when both letters were legal. The observed equivalence between these latter two cases is critical evidence against higher-order nonindependence models. Subjects should have done much better in the both legal than the both illegal contexts according to these models. This follows from the fact that the context modifies lower-level analyses.

The independence model provides a good description of the results. The root mean square (RMS) deviation was .02. It is clear that an additive combination of the two sources would fail since the curves would have to be parallel (see Massaro & Cohen, 1976). Supporting this, the RMS was 5 times larger for the additive than the multiplicative model. The results, therefore, provide strong support for the idea that orthographic structure and visual feature information provide independent sources of information in reading. Any assumption of higher-order analyses overriding and changing the nature of lower-level analyses is unwarranted. Until an equally parsimonious description of the results can be given by nonindependence views, we must accept the independence model.

Descriptions of Orthographic Structure

The explanation offered for enhanced recognition of a word is based on the idea that the reader actively utilizes the orthographic structure to aid perceptual processing. Given this interpretation the question that becomes critical is, "What is the psychologically relevant measure of orthographic structure?" There appear to be two major approaches to the description of this structure (Venezky & Massaro, 1976). In the probablistic approach, word tokens sampled from text are used to define the probability of occurrence for single letters, bigrams, and so on. These data allow generation of letter strings that differ in their approximation of English. In the rule-governed approach, legal strings are generated based on generalizations about the underlying patterns of English orthography. Rules supposedly define which letters or letter sequences are allowed (or not allowed) in which positions or graphemic context. The patterns of the orthography follow from scribal practices and phonological constraints of the language. One critical difference between the two approaches is that the probabilistic description is fundamentally dependent on the frequency with which letter sequences occur in English whereas the rule-governed description may generate a "regular" sequence of letters that does not occur in real words and may reject some sequences that do occur.

Massaro, Venezky, and Taylor (in press) evaluated the vali-

dity of probabilistic and rule-governed descriptions as psychologically real measures of intra-word redundancy. Six-letter strings were generated by rearranging the letters of six-letter words to give strings that were independently high or low in intra-string redundancy according to the probabilistic and rule-governed approaches. A probable or improbable string by the probabilistic description could be either high or low in orthographic regularity by the rule-governed description. By independently varying strings along both of these dimensions, we hoped to assess the degree to which these descriptions were psychologically real measures of intra-word redundancy.

The probabilistic description of "positional frequency" was used to generate strings that were high or low on this dimension. The positional frequency of a given letter in a letter string is measured by the frequency with which the letter occurs in the same spatial position in words of the same length. The positional frequency of a letter string is the sum of the positional frequencies of each of the individual letters in the string. The positional frequencies were computed from the tables of Mayzner and Tresselt (1965) which gave the total number of occurrences of each letter at each spatial pattern in a sample of 20,000 words of running test.

The letter strings were also varied on the dimension of orthographic regularity as defined by rules developed by Venezky and Massaro (1976). The rules are based on graphemic conventions and phonological constraints (Venezky, in press; Venezky and Massaro, 1976). Graphemic conventions can usually be traced to scribal practices, printing conventions, and lexicographer's selections. In English words, for example, no letter can triple and only 16 letters double. Doubled letters are not permitted in initial word positions. Some single letters also have positional constraints, such as j, u, and v do not occur in final position. It should be noted that regularity as defined by graphemic convention is not equivalent to pronounceability. For example, lev and baaaf are highly pronounceable but irregular. The first, lev, has v in final position and the second, baaaf, has a tripled letter.

Phonological constraints provide the second source of restrictions on letter sequences in English words. For example, certain sound sequences are not permitted in English words. Some consonant sequences do not occur in initial position, such as ds, a stop consonant followed by a fricative. Others do not occur in word final position, such as sd, fd, pg, and so on.

The letter strings selected for the experiments allowed us to evaluate the degree to which summed positional frequency and our rule-governed description are psychologically relevant measures of intra-word redundancy. We began by selecting letter strings that were high or low in summed positional frequency, and then chose strings in both of these categories which were

also orthographically regular or irregular. Given the word remain, the permutations naimer and aeminr are both high in summed positional frequency but the first string is orthographically regular whereas the second is not. The strings ineram and rmenai are the analogous regular and irregular low frequency strings.

In one task, the subject was presented with one of the test strings followed immediately by a masking stimulus. A target letter followed and the subject indicated whether or not the target letter was present in the test string. The results indicated that both orthographic regularity and summed positional frequency facilitated recognition of the letters in the test string. Strings high in orthographic regularity were recognized 4 percent more accurately that irregular strings. Strings high in summed positional frequency were recognized 5 percent more accurately than strings low in positional frequency. There was no interaction of these two variables. This initial work and more recent studies (Taylor, Massaro, and Venezky, 1977) make it apparent that both rule-governed and frequency-based descriptions will be necessary to capture the reader's knowledge of orthographic structure.

Semantic vs. Orthographic Context

To what extent does the word advantage result from a lexical representation for words but not for nonwords or most single letters? One method of answering this question has been to compare recognition of words to recognition of pseudowords which are equivalent to words in all respects except wordness. In contrast to Baron and Thurston's (1973) finding, Juola, Choe, and Leavitt (1974) and Manelis (1974) found significant differences between real words and pseudowords in a Reicher (1969) task. Overall, a letter in a word was correctly reported about 5 percent more often than a letter in a pseudo-word. The results are not necessarily inconsistent with a strict orthographic structure interpretation, however.

Demonstrating that words are seen better than pseudo-words does not necessarily mean that lexical assess contributes to the visual resolution of the letter string. That is to say, words may be recognized better than pseudo-words not because there is a lexical representation for words but not pseudo-words, but simply because words are more orthographically regular than pseudowords. It is very difficult to control exactly for orthographic regularity in word pseudo-word comparisons. For example, Manelis changed only one letter of a word to make it a pseudo-word, but, in some cases, the yoked pseudo-words appeared to be much less regular than the parent words (compare west and town to weft and towd, respectively). McClelland (1976) equated the words and pseudo-words on position specific single-letter and bigram frequency but Massaro, Venezky, and Taylor (in press)

have shown that this measure is not a sufficient description of orthographic structure.

A word advantage over pseudo-words may also result from differential forgetting in short-term memory rather than differences in letter recognition. Given a short test display, followed immediately by a masking stimulus, the subject has a very weak visual code and must translate this into a more abstract code in order to choose between the test alternatives. Let us assume that subjects depend on both a phonemic and semantic rendering of their visual experience. As an example, consider the comparison between the word heat and the pseudo-word veat in the case in which both items were correctly recognized and encoded. The word will have the advantage of a semantic code, if some of the phonemic code is lost. If the subject forgets a single phonetic feature of the final consonant, such as place of articulation, he or she will not be able to choose between the test pair veat and veap in the pseudoword case. This forgetting isn't a problem in the word case because the semantic code will allow him or her to choose heat given the alternatives heat and heap.

Another advantage of words over pseudo-words may be due to the nature of the regularity rules that the reader has internalized. Readers may base their rules primarily on words that do occur in the language rather than on words that can occur. Independently of lexical representations, readers may know which letters in particular contexts spell words. For example, consider the orthographic context wai preceding a fourth letter in a four-letter string. It is clearly the case that both d and t are orthographically regular in this position, and therefore, either waid or wait are good candidates for English words. Given the fact that only wait is a word, the reader may have stored orthographic information indicating that t is to be preferred over d in the context wai. Faced with this context in reading, then, the reader will synthesize t rather than d if the visual information has narrowed down the final letter to these two alternatives. Given the pseudo-word context cai, on the other hand, both d and t are orthographically regular and neither of these are words so that the reader should not prefer one letter over the other. If this is the case, the reader will more accurately process a t in the context wai than in the context cai. This analysis of an advantage of a word over a pseudo-word does not assume that lexical access is necessary. It is conceivable that a reader could have a brain lesion that would preclude lexical access of wait. This same reader would choose wait as a word in a choice between waid and wait, even though he is unable to give the meaning of wait. This same subject could, therefore, show an advantage of words over pseudo-words in the target search task without any lexical access.

Controlling for Orthographic Redundancy

In the work of Thompson and Massaro (1973) and Massaro (1973) it was suggested that sufficient practice with a fixed set of alternatives would preclude the utilization of redundancy in word strings. To test this idea, the subjects were tested with letters P, R, C, and G, with the words APE, ARE, ACE, and AGE, and the nonwords VPH, VRH, VCH, and VGH. If utilizing orthographic structure is responsible for the word advantage in the post-cue Reicher paradigm, then words should no longer show a processing advantage when the alternatives are specified in advance. Supporting this, letters presented alone were recognized better than words or nonwords which did not differ. Bjork and Estes (1973), Estes (1975), and Estes, Bjork, and Skaar (1974) have found similar results when the fixed set of alternatives R and L are used throughout the experiment.

More recently, data have been presented that demonstrate that a word advantage can be shown with target alternatives given before the test display. Carr, Lehmkuhle, Kottas, Astor-Stetson, and Arnold (1976) extended the previous work with the response alternatives P, R, C, and G, but with a stimulus set that permitted the target to occur at any of the three display locations. Performance was significantly better for words when the test letter occurred in the first or third position. Spector and Purcell (1977) replicated the paradigm of Estes and his colleagues and showed a word advantage over nonwords with the restricted set of alternatives R and L. These two sets of results show that a word advantage can be found with prior knowledge of the alternatives, when the position of the test letter is uncertain. In terms of our analysis, prior knowledge of the alternatives does not always ensure that orthographic structure of letter strings will not be exploited. It is up to the reader to decide whether he or she will restructure the orthographic rules that have been used for a lifetime of reading. In many cases, it will be easier to persist in normal reading even though the experimenter has defined a limited set of alternatives. It is unfortunate, but the results reflect the all too-common finding that an experimental manipulation cannot be expected to determine exactly which psychological process will be operational in the task.

Other Evidence

Pollatsek and Carr (in press) have provided a very careful review of work in same-different matching tasks, a domain not addressed in Massaro (1975a) or in this review. What is encouraging from their review is that the authors have arrived at very similar conclusions from an entirely different data base. Meaning and supraletter features in words cannot account for the

word advantage in same-different tasks, as it cannot account for the word advantage in the tasks discussed here. Similarly, frequency and familiarity based descriptions of orthographic structure do not appear to be sufficient to describe the knowledge readers bring to bear on the perception of letter strings. Converging viewpoints such as these hopefully mean that we are reducing uncertainty about how language is perceived.

IV

SECONDARY RECOGNITION

Secondary recognition transforms synthesized percepts into meaningful forms in generated abstract memory. In speech perception, it is assumed that the input is analyzed syllable by syllable for meaning. In reading, letter sequences are closed off in word units. In both cases, the secondary recognition process makes the transformation from percept to meaning by finding the best match between the perceptual information and the lexicon in long-term memory. Each word in the lexicon contains both perceptual and conceptual information. The concept recognized is a function of two independent sources of information: the perceptual information in synthesized memory and the syntactic/semantic context in the message.

PERCEPTUAL AND CONTEXTUAL CONTRIBUTIONS TO LISTENING

Abstracting meaning is a joint function of the independent contributions of the perceptual and contextual information available. In one experiment, Cole (1973) asked subjects to push a button every time they heard a mispronunciation in a spoken rendering of Lewis Carroll's Through the Looking Glass. A mispronunciation involved changing a phoneme by 1, 2, or 4 distinctive features (for example, confusion mispronounced as gunfusion, bunfusion, and sunfusion, respectively). The probability of recognizing a mispronunciation increased from 30 percent to 75 percent with increases in the number of feature changes, which makes apparent the contribution of the perceptual information passed on by the primary recognition process. The contribution of contextual information should work against the recognition of a mispronunciation since context would support a correct rendering of the mispronounced word. In support of this idea, all mispronunciations were correctly recognized when the syllables were isolated and removed from the passage.

Cole and Jakimik (1977) extended Cole's (1973) mispronunciation task to evaluate how higher-order contextual information can influence sentence processing. To the extent that a word is predicted by its preceding context, the listener should be faster at detecting a mispronunciation. This follows from the idea that the quickest way to detect a mispronunciation is to first determine what the intended word is and then notice a mismatch with what was said. Given the sentences, "He sat reading a book/bill

until it was time to go home for his tea," mispronouncing the /b/ in book as /v/ should be detected faster than the same mispronunciation of bill. In fact, listeners were 150 msec faster detecting mispronunciations in highly predictable words than in unpredictable words.

In other experiments Cole and Jakimik (1977) demonstrated similar effects of logical implication. Consider the test sentence, "It was the middle of the next day before the killer was caught," with the /k/ in killer mispronounced as /g/. Detection of the mispronunciation should be faster when the test word is implied by the preceding sentence, "It was a stormy night when the phonetician was murdered," compared to the case in which the preceding sentence states that the phonetician merely died. Thematic organization also facilitated recognition of words in their stories. Given an ambiguous story, a disambiguating picture shortened reaction times to mispronunciations of thematically related words but not to mispronunciations of other words that were unrelated to the theme of the story.

Marslen-Wilson (1973) asked subjects to shadow (repeat back) prose as quickly as they heard it. Some individuals were able to shadow the speech at extremely close delays with lags of 250 msec, about the duration of a syllable or so. One might argue that the shadowing response was simply a sound to sound mapping without any higher order semantic-syntactic analyses. When subjects make errors in shadowing, however, the errors are syntactically and semantically appropriate given the preceding context. For example, given the sentence "He had heard at the Brigade," some subjects repeated "He had heard that the Brigade." The nature of the errors did not vary with their latecy; the shadowing errors were always well-formed given the preceding context.

In another experiment (Marslen-Wilson, 1975), subjects shadowed sentences that had one of the syllables mispronounced in a three-syllable word. Subjects never restored the word, that is, repeated back what should have been said when the mispronunciation occurred in the first syllable. With mispronunciations in the second and third syllables, however, a significant proportion of restorations occurred. If the mispronounced word was also syntactically and semantically anomalous, restorations did not occur for any mispronounced syllable. These results indicate that restorations will not occur if the shadower does not have sufficient acoustic information and syntactic/semantic context to make the restoration appropriate. If the context were the overriding factor, subjects should have replaced the syntactically/semantically anomalous word with the appropriate word. This, however, did not occur.

Marslen-Wilson and Tyler (1975) has subjects monitor sentences for a target item in three types of target monitoring tasks. The target item was either a particular word, any word

that rhymed with the target, or a member of superordinate category. Three types of sentences were used: (1) normal, (2) syntactically correct but semantically anomalous by randomization of the content words, and (3) a completely random ordering of words in the sentence. The mean reaction times for detecting the target was a function of both the monitoring task and the sentence structure. Sentence structure facilitated monitoring in all three tasks showing that higher-order constraints were functional, regardless of the target analyses that were required. The reaction times were shortest for detecting a specific target, next shortest for a rhyme, and longest for a member of a superordinate category.

Marslen-Wilson and Welsh (in press) asked observers to shadow spoken passages from a popular novel. The words of the passage were read to the subjects at a rate of 160 words per minute. The subjects were told to repeat back exactly what they heard. At random throughout the passage, common three-syllable words were mispronounced. When the words were mispronounced, only a single consonant phoneme was changed to a new consonant phoneme. The new phoneme differed from the original by one or three phonemic distinctive features, based on Keyser and Halle's (1968) classification system. Independently of the degree of feature change the changes could occur in the first or third syllable of the three-syllable word. Finally, the mispronounced words were either highly predictable or unpredictable given the preceding portion of the passage. Subjects were not told that words could be mispronounced although they probably became aware of this early in the experiment. All subjects shadowed at relatively long delays greater than 600 msec. The primary dependent measure in the task was the percentage of fluent restorations, that is, the proportion of times the shadowers repeated what should have been said rather than what was said. About half of the mispronounced words were restored, the restorations were made on-line with an average latency, and the shadowing was not disrupted. (When the mispronunciation was repeated exactly, i.e., not restored, shadowing was disrupted and response times increased.)

The change in the percentage of restorations as a function of the three independent variables in Marslen-Wilson and Welsh's study can illuminate how the listener integrates acoustic information and higher-order context. Figure 6 presents the observed results in terms of the percentage of fluent restorations. All three variables influenced the likelihood of a restoration: shadowers were more likely to restore a one-feature than a three-feature change, a change in the third than the first syllable, and a change in a highly predictable than an unpredictable word.

Marslen-Wilson and his colleagues interpret this series of experiments as evidence against serial theories of language processing, which assume that "varying degrees of delay before information at any one level of analysis can interact with information at a higher level" [Marslen-Wilson & Tyler, 1975, p. 784].

However, the results do show exactly such a delay. Restorations seldom occur when the first syllable is mispronounced by 3 features even though the word is relatively probable given the preceding context. This means that some low-level perceptual analyses of the word occurred regardless of the higher-order constraints and then the outcome of this analysis was combined with the higher-order constraints. The fact that higher-order constraints in the passage influence shadowing does not mean that some analyses do not begin before others. More importantly, their view might be interpreted to mean that higher-order analyses modify the output of lower level analyses. However, a quantitative model that assumes that both levels of analyses are functionally independent can accurately describe the results of their experiments. Figure 6 also presents

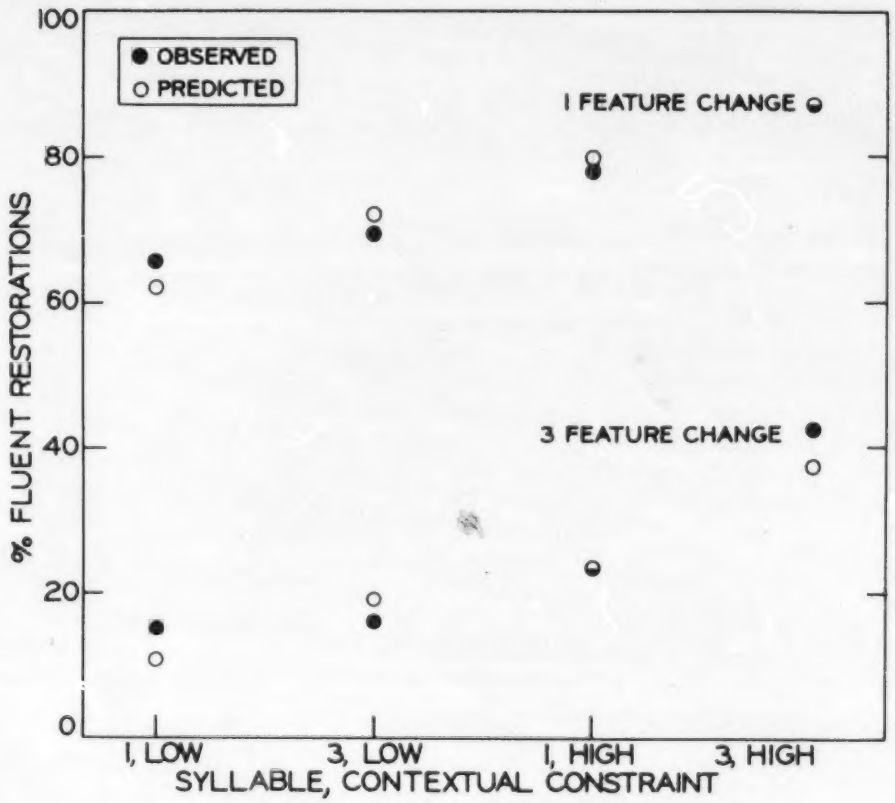

Figure 6. Predicted and observed percentage of fluent restorations as a function of the amount of feature change, the syllable, and the contextual constraint of the mispronunciation (after Marslen-Wilson & Welsh, in press).

the predictions of a quantitative formulation of the independence model.

The model assumes that the information passed on by the feature detection and evaluation process is equivalent regardless of the higher order constraints in the message. The listener should be more likely to restore a single-feature change than a three-feature change. Substituting /f/ for /v/ should be more confusable than substituting /n/ for /v/. That is to say, /f/ gives the subject more acoustic information for /v/ than does /n/. Letting a_i represent the acoustic information available for the original sound, then a_1 for a one feature change should be larger than a_3 for a three-feature change. Mispronouncing the word in the first or third syllable should also influence the likelihood of a restoration. I have argued previously that speech is processed syllable by syllable (Massaro, 1975a) and the listener attempts to make lexical access at each syllable boundary. We might expect that lexical access of an upcoming word is usually not possible before some of the word is heard. However, given the first two syllables of a three-syllable word, the listener has a good chance at recognizing the word correctly. This means that the lexical information available for the mispronounced word should be much higher at the time of the third than the first syllable. Accordingly, a subject should be much more likely to restore a mispronunciation in the third than the first syllable. Letting ℓ_i equal the amount of lexical information that is available, ℓ_3 should be much larger than ℓ_1, where 1 and 3 refer to a mispronunciation in the first and third syllables, respectively. Finally, highly constrained context should make a restoration much more likely than unconstrained context. Therefore, letting s_c and s_u represent the amount of preceding sentential information under constrained and unconstrained contexts, we would expect s_c to be much larger than s_u.

The critical feature of our analysis is that each of the three sources of information--acoustic, lexical, and sentential--are assumed to have independent influences on a restoration in this example. That is to say, in contrast to some top-down processing models, highly constrained context does not change low-level perceptual analyses. In this model, the values of a_1 and a_3 do not change as a function of the amount of preceding sentential context. Similarly, low-level perceptual processes do not override higher-order constraints: the values of s_c and s_u do not change as a function of the acoustic information that is available.

In order to predict the probability of a restoration it is necessary to define how the three sources of information are combined and integrated into a synthesized language percept. In our analysis, we classify a shadowing response as either a restoration, R, or a nonrestoration $\bar{R}$. It is assumed that the probability of a restoration, P(R) is a function of three factors: the acoustic information available, the syllable that is mispronounced, and the amount of preceding context.

$$P(R) = f(a \wedge \ell \wedge s) \quad (3)$$

What is critical is how conjunction is defined in this model. Massaro and Cohen (1976) and Oden and Massaro (1977) have shown that a multiplicative combination provides a good description of the combination of acoustic features in speech perception. If the model is extended to this situation, then the degree, Degree (R), to which the speech sound will be heard as what should have been said can be described by:

$$\text{Degree } (R) = a * \ell * s \quad (4)$$

where a, ℓ, and s are as defined earlier. To predict the likelihood of a restoration, it is necessary to determine the degree that the sound is heard as it should have been said relative to the sound being heard as anything else. Hearing the sound as anything else is the negation of hearing the sound as what should have been said. In this case, if a represents the amount of acoustic information for a restoration, the value (1-a) represents the amount of information for a nonrestoration. Quantifying this prediction, we get:

$$P(R) = \frac{a\ell s}{a\ell s + (1-a)(1-\ell)(1-s)} \quad (5)$$

In the analysis under consideration, there were two levels of acoustic information, two levels of lexical information, and two levels of sentential information. Therefore, it is necessary to estimate two values each for a, ℓ, and s, in order to predict performance under the eight experimental conditions. Figure 5 presents the observed and predicted results. In addition to the good fit of the model, the parameter estimates confirm our analysis. The value of a for a single-feature change was .76 whereas it was only .24 for a three-feature change. The value of ℓ representing the degree of lexical constraint was .44 when the first syllable was mispronounced and .56 when the third syllable was mispronounced. Finally, the degree of sentential constraint was .60 with highly predictable words and only .38 with a word nonpredictable from the preceding context.

These results support the common belief that both low-level perceptual information and high-level contextual information contribute to speech processing (and reading). What remains to be discovered is exactly how these two sources of information are combined in the abstraction of the meaning of the message. Empirically, this calls for the independent variation of the two sources of information. On the theoretical level, it is necessary to develop and test quantitative models of how the sources of

information are integrated into meaning. It seems likely that many of the upcoming generation of language experiments will be addressed to this issue.

PHONOLOGICAL MEDIATION

One persistent question in reading-related research is the extent to which the reader translates print into some kind of speech code before meaning is accessed. A similar but not identical question is the extent to which the speech code is necessary for the derivation of meaning. Figure 7 presents two extreme answers to the phonological mediation question. In the first model, letters are identified and mapped into a speech code using spelling-to-sound rules, and meaning is determined on the basis of the derived speech code. In the second model, meaning is determined from letter resolution and a speech code is not made available until after meaning has been accessed. Massaro (1975a) reviewed this issue in some detail, covering recent theories of phonological mediation and relevant experimental data. He concluded that, "although phonological mediation can be rejected in word recognition, it may play a significant role in later pro-

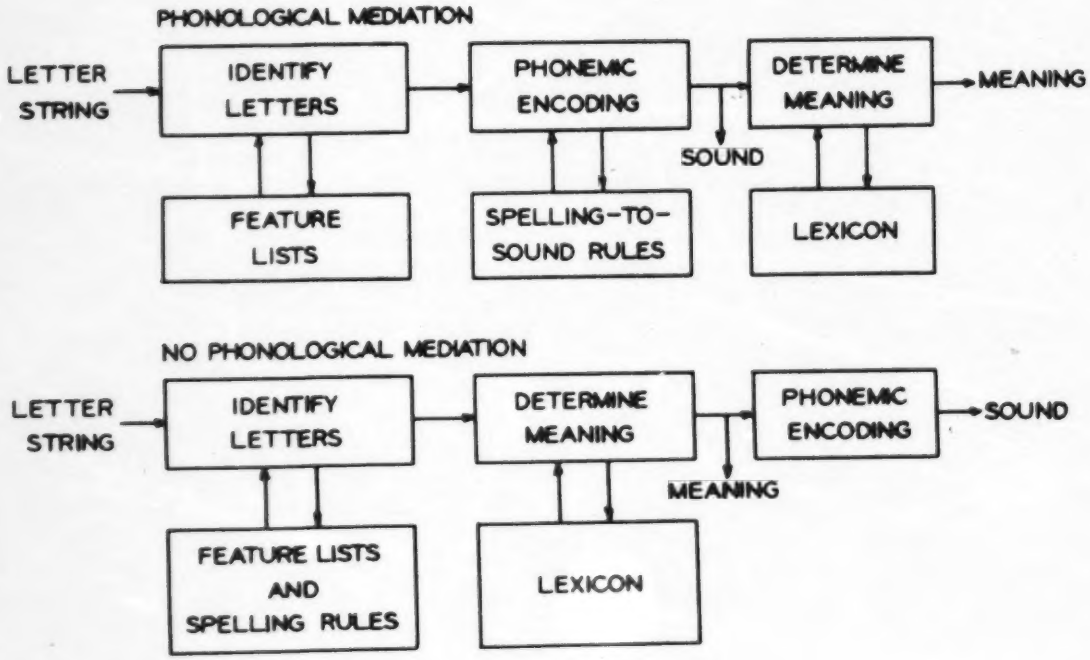

Figure 7. Two stage models of the role of phonological mediation in reading.

cessing stages such as rehearsal and recoding in generated abstract memory [pg. 273]." Although a relatively large amount of research has been added to the data base, the story does not require any modification.

Gough and Cosky (1976) believe that they have accumulated some new data in support of the phonological-mediation view of Gough (1972). Subjects were asked to read aloud as rapidly as possible words that violated or obeyed spelling-to-sound rules. If phonological mediation occurs, regular words which conform to spelling-to-sound rules should be converted to a speech code faster than exception words which violate the rules. Accordingly, the time to comprehend the word and name it aloud should take longer for the words that violate spelling-to-sound rules. In support of their hypothesis, the pronunciation times for exception words averaged 27 msec longer than the pronunciation times for regular words. However, there is no assurance that differences in pronunciation time result from differences in word recognition time. The differences in reaction time could also have resulted from differences in the time for response selection and programming after the word had already been identified (see Massaro, 1975a, p. 262).

In order to provide evidence that differences in naming times directly reflect differences in word recognition time, it is necessary to perform a stage analysis of the naming task and to include a number of other independent variables known to influence specific stages in the task (Massaro, 1975b; Sternberg, 1969). Consider this task in terms of the model depicted in Figure 1. Naming a written pattern requires feature detection, primary recognition, and secondary recognition processes and also response selection and response programming operations. In terms of this analysis the RT between the onset of the written pattern and the onset of the spoken response is a composite of 5 component times:

$$RT(\text{name}) = FD + PR + SR + RS + RP \qquad (6)$$

where FD, PR, SR, RS, and RP represent the times for the five respective processes.

The critical components in the naming task include the time to resolve the shape of the pattern (that is, to see the pattern), attach a name (speech code) to the seen pattern, to select the appropriate articulatory program for the speech code, and finally to program the articulators to execute the response. Response execution time would not contribute to the actual RT since the RT is measured at the onset of the naming response. One simplification of the analysis would be to divide up the RT into input and output operations. In this case, the first three operations would entail stimulus processing, whereas the last two would represent response operations.

$$RT = SP + RO \qquad (7)$$

where SP and RO equal the times for stimulus processing and response operations, respectively.

The present concern for localizing naming-time differences at a particular stage of processing is not new; in fact, James McKeen Cattell (1888) provided exactly this analysis almost 100 years ago. Many contemporary investigators have cited Cattell's finding that a short word can be named in less time than a single letter, with the implication that words are perceived in less time than are single letters. Cattell realized, however, that the naming task included both perception time and "will-time," as he called the time to choose a response. Relative differences in perception time were determined by using a Donders Type C reaction, in which the observer was required to make a simple response such as lifting a finger off a key in response to just one of many possible alternative stimuli. In different tests, the subject was asked to respond to just one of many possible letters, words, colors, and so on. In contrast to the naming task, the results showed shorter reaction times for single letter than word alternatives. Cattell's interpretation appears to be still valid today. To quote from his Popular Science Monthly article, in 1888, The time it takes to think, "A letter can be seen more quickly than a word, but we are so used to reading aloud that the process has become quite automatic, and a word can be read with greater ease and in less time than a letter can be named [p. 23, my italics]." In a similar set of studies, Cattell found that it takes less time to perceive color than a letter or word but much longer to name the color relative to naming the letter or word. Huey (1908) also believed that letter recognition mediated word recognition, but that the context set for words "drafts to itself the energy which would have been given to the letters, [p. 113]." This is analogous to the Stroop color-word phenomenon.

That naming RTs depend on response selection time is apparent in a series of experiments showing that the RTs increase with increases in the number of syllables of the word to be named (Eriksen, Pollack, & Montague, 1970; Klapp, 1974). Klapp, Anderson, and Berrian (1973) attributed the naming differences to response selection and preparation since similar tasks requiring recognition without pronunciation showed no effects of syllable length. In addition, picture naming was found to be syllable-dependent in the same way as words. Frederiksen and Kroll (1976) found large effects of word length and syllable structure on naming RTs but no effects of these variables on the time to make lexical decisions. These results make it apparent that response must be accounted for in naming studies of word recognition.

Theios and Muise (1976) compared pronunciation times for real words and pseudowords, while attempting to control for the articulation response. For each real word, there was a yoked

pseudoword which was homophonic to the word, but spelled differently. Given that articulation differences between words and their corresponding pseudowords would be eliminated, the pronunciation times might provide a more direct index of recognition times. Naming times were 20 msec longer for pseudowords than for real words. Frederiksen (1976) carried out a similar experiment, but created his pseudowords by changing a single vowel in each of the real words. However, rather than randomizing words and pseudowords in the same session as in the Theios and Muise study, words were tested in one session and pseudowords in another. The pseudowords took 142 and 49 msec longer to pronounce for poor and good high-school readers, respectively.

The advantage of words over pseudowords can be interpreted to mean that subjects can more quickly recognize a word and retrieve the appropriate articulatory program stored in lexical memory than they can map a nonword letter string into a phonological code based on spelling-to-sound rules of the language. If derivation of a phonological code always preceded lexical assess, as assumed by Gough (1972), then naming times should have been equivalent for words and pseudowords. Given that words and nonwords were equated on phonology in the Theios and Muise study and closely matched on orthography in the Frederiksen study, it seems likely that the word advantage over pseudowords reflects processes that depend on lexical access for words but not pseudowords. Frederiksen may have found much larger differences than Theios because blocking the words in a session would encourage the subjects to pronounce the words via lexical access. Randomizing words and pseudowords in the Theios and Muise study might have encouraged pronouncing some of the words by way of spelling-to-sound rules rather than by way of lexical access. In agreement with this interpretation, Frederiksen and Kroll (1976) found a larger effect of word frequency on naming RTs when only words were presented in a block of trials relative to a random mixture of word and pseudoword trials.

Green and Shallice (1976) asked subjects to judge whether two words rhymed or whether they belonged to the same broad semantic category. Misspelling the words as homophones produced a much larger decrement in the semantic than the rhyming task. If lexical access occurs via phonological coding, there is no reason that the semantic task should have been slowed more by misspelling than the rhyming task was. The fact that the rhyming task was performed about twice as fast as the semantic task shows that lexical access was not necessary in the former task although it was in the latter. Spelling-to-sound rules would have been sufficient to perform the rhyming task and misspelling should have very little affect on this process. In support of this, misspelling the words increased reaction times by only 11 percent. Lexical access should be drastically influenced by misspelling however, if it occurs via a visual code. Reaction times were slowed by 58 percent in the semantic task, arguing

against the idea of phonological or speech recoding in lexical access and derivation of meaning. The results support other negative findings on the necessity of phonemic encoding in processing written language for meaning (see Massaro, 1975a).

VISUAL, PHONOLOGICAL, AND SEMANTIC CODES IN READING

Regardless of how letters, letter strings, and words are resolved visually, there is growing evidence that this visual information must be transformed into another code very quickly after stimulus presentation. Observers in tachistoscopic tasks appear to remember very little about the visual properties of the short display presentation. Thompson and Massaro (1973) and Juola et al. (1974)varied the similarity of the two forced-choice alternatives in the Reicher task. Subjects saw a word or letter followed by two letter alternatives and reported which letter was present in the display. Performance did not differ for confusable and distinctive letter alternatives. This result indicates that observers were not able to maintain partial visual featural information to aid in their choice between the letter alternatives. The subjects must have recoded the information into whole letters or words or an abstract code, preventing the utilization of partial visual information in this task (see Massaro, 1975a).

McClelland (1976) presented words and letter strings in upper, lower, or mixed cases. In addition to reporting letters themselves, McClelland asked the subjects to indicate whether the letter strings were upper, lower, or mixed case. Subjects averaged only 39 percent correct where chance would be 33 percent. Even when subjects were able to report all of the four letters in the string in a whole report, they were able to identify the appropriate case only half of the time.

The fact that the visual information is so brief and fleeting in a tachistoscopic presentation makes apparent the need for recoding what has been resolved in the short glance. Mezrich (1973) was the first to pursue recoding differences as an explanation of the word advantage over single letters. Pilot subjects had reported encoding the word stimuli as complete words and the letter stimuli as something less than unitary wholes. If verbal encoding of words is responsible for the word advantage, then the word advantage should disappear when letters are also verbally encoded. To this end, he required the subjects to vocalize each stimulus presentation even if they were not sure of what to say. Then after 1.5 sec the test alternatives were presented. In this study, a letter presented alone was recognized slightly better (6 percent) than a letter presented in a word. In a control study without vocalization, a 6 percent word advantage was found. Unfortunately, there were a number of methodological problems in making this comparison; for example, the time between the onset of the test display and the choice alternatives

was longer in the vocalization than in the control condition and there is some evidence that a delay works against the word advantage (see Wheeler, 1970).

Hawkins, Reicher, Rogers, and Peterson (1976) evaluated the degree to which subjects are dependent on a phonological code in the Reicher paradigm. On word trials, the two alternatives were either phonetically identical (cent, sent) or different (sold, cold). Observers were given either a large or small proportion of word trials that had homophone alternatives. Single letter trials were also interspersed throughout the session. In the condition with a high proportion of homophones, Hawkins et al. found a 15 percent advantage over single letters. More importantly, the size of the word advantage was the same for trials with homophone and nonhomophone alternatives. Accordingly, a phonetic code did not improve performance in this task either because subjects did not use it or because it did not contribute additional information beyond that given by other (semantic or visual) codes.

In the condition with a low proportion of trials with homophone alternatives, however, the word advantage was only 6 percent with homophone alternatives and 20 percent with nonhomophone alternatives. It is clear that subjects relied on a phonetic code in this condition and this code improved performance 14 percent relative to information given by other codes. When the homophone alternatives were presented on only a small proportion of trials, subjects did less well than on nonhomophone trials. With a large proportion of homophone-alternative trials, however, there was no difference between the two kinds of trials. The authors argue that phonetic coding is usually used in this task, but another code can be used if the phonetic code is shown not to distinguish between the alternatives in the task. It would be interesting to test whether the alternative code was visual or semantic. One possibility would be to employ pseudowords rather than real words to eliminate the possibility of a semantic code. If the same results are obtained, then it would appear that a visual code can be used. If subjects do poorly with homophone alternatives even when they expect a high proportion of homophone alternatives, however, the results would say that a lasting visual code is not available in tachistoscopic experiments.

The Hawkins et al. study cannot be taken as evidence that speech encoding mediates word recognition. There is no measure of the temporal course of processing; therefore, we do not know whether the phonemic code was derived before or after a semantic one. Why might a phonemic code have been more successful than a visual or semantic code in Hawkins' study? One reason is that subjects may have had both the phonemic and semantic codes available but they were more likely to discard or forget the semantic code, especially when the phonemic code was usually

successful in the low-homophone condition. A second reason that the phonemic code may have been more successful than a semantic code is that subjects were unlikely to recognize the whole word in the task. If correct recognition of a single letters was .7, then recognition of all four letters would be $.7^4$ or about one time in four. Given partial information about the letter string and the fragility of a visual representation, subjects might be likely to utilize a speech code in the task. Given a phonemic code in the low homophone condition, they would be more likely to be correct when the word alternatives were phonemically different. The fact that homophone alternatives do not disrupt performance in the high-homophone condition shows that other codes besides a speech code are available in the task.

V

REHEARSAL AND RECODING

In this model, the same abstract structure stores the meaning of both listening and reading. Generated abstract memory (GAM) in our model corresponds to the working memory of contemporary information processing theory. Rehearsal and recoding processes operate at this stage to maintain and build semantic/syntactic structures. There is good evidence that this memory has a limited capacity, holding about 5 ± 2 chunks of information. For a more detailed discussion of processing at this stage, see Massaro (1975b, Chapter 27). In this section, we review some evidence that the surface structure of GAM in reading and listening is organized in terms of speech codes.

SPEECH CODES IN GENERATED ABSTRACT MEMORY

Baron (1973) studied the role of phonemic encoding in processing short phrases for meaning. Subjects were asked to indicate whether or not short phrases made sense either according to their spelling or their sound. Examples were He Is Ill, Tie the Not, and Ill him for the phrases that respectively make sense visually and phonemically, just phonemically, and do not make sense. If subjects encode the phrases in a speech form, they should be disrupted in the spelling test when shown sentences like Tie the Not since it makes sense in sound but not in the spelled form. Although subjects were able to reject the homophone phrases like Tie the Not just as quickly as the nonsense phrase like Ill him in the spelling task, they made significantly more errors with the homophone phrases. Although this result shows that a speech code may be functional in the spelling task, it may have come into play after lexical access and the derivation of meaning (Massaro, 1975a, pg. 266-271). In the sound task, subjects took longer and made more errors on the homophonically correct phrases than the visually correct phrases. This result indicates that meaning was more readily derived directly from print than sound in this task.

Kleiman (1975) has recently presented a nice series of experiments showing that although speech recoding is not necessary in deriving meaning from print, it is a critical determinant of storage in working memory. The work centers on the idea that the disruptive effect of concurrent shadowing in different tasks reveals the degree to which speech recoding occurs in each of the

tasks. Subjects performed each of three tasks in the first experiment. Subjects saw two words side by side and had to indicate whether the words were graphemically similar, phonemically similar, or semantically similar. For example, the subject in the phonemic task would respond true to tickle-pickle but false to lemon-demon since the former rhyme but the latter do not. A "false" response in the graphemic task was made if the words differed by more than a single letter. Each of the three tasks was performed with and without the shadowing task. It is unlikely that the graphemic task requires speech decoding and the interference caused by shadowing should provide an index of the general disruptive effects of shadowing. The additional interference caused by shadowing in the phonemic task, which requires speech decoding, indexes the direct interference of shadowing on speech decoding. Given these measures, the degree of interference caused by shadowing on the semantic task should reveal the degree that speech decoding is present in this task. The results showed that the interference caused by shadowing was equivalent in the graphemic and semantic tasks and significantly greater in the phonemic task. In terms of the logic of this study, this means that the amount of phonemic encoding was equivalent in the graphemic and semantic tasks and significantly greater in the phonemic task. In a second experiment Kleiman showed that phonemic similarity did not disrupt the graphemic task, which shows that no phonemic encoding occurred in the graphemic task. Therefore, phonemic encoding did not occur in the semantic task either since shadowing had equivalent effects in the graphemic and semantic tasks.

In a third experiment, observers were given five-word sentences in two different tasks with and without shadowing. In the category task, subjects were given a category name and asked to indicate whether or not the following sentence contained a member of that category. In the acceptability task, the subject indicated whether or not the word string was a semantically acceptable sentence. Shadowing increased reaction time 78 msec in the category task and 394 msec in the acceptability task. Given that the category task could be performed without much working memory load whereas the acceptability task could not, the author concludes that speech recoding is important in storage in working memory but not in initial processing for meaning. Accordingly, it appears that speech recoding is critical in reading text for meaning but at a later stage than word recognition. The result is consistent with the model shown in Figure 1 where both auditory and visual processing converge at the level of generated abstract memory.

Erickson, Mattingly, and Turvey (1974) present some evidence that GAM for Japanese Kanji is also based on a speech code. Their Japanese subjects had more difficulty in immediate memory for phonetically similar and homophonic words than for semanti-

cally similar or unrelated words. These findings are all the more impressive because Japanese Kanji are logographic and supposedly do not necessitate phonetic encoding. Given the short-term memory requirements, however, the phonetic reading of the Kanji was a better medium for retention than the corresponding meaning from which the phonetic reading may have been derived.

VI

CONCLUSION

It seems valuable to attack reading and listening with similar methodological and theoretical forces in the framework of an information-processing model. Our concern is with how the reader and listener perform, and with the dynamics of this performance. Although the surface structures of written text and speech present questions unique to each skill, the apparent similarities in deep structure offer the hope of a single framework for understanding both reading and listening.

REFERENCES

Allport, D.A. On knowing the meaning of words we are unable to report: The effects of visual masking. In S. Dornic (Ed.) Attentional Performance VI, Hillsdale, N.J.: Erlbaum, 1977.

Anderson, N.H. Information integration theory: A brief survey. In D.H. Krantz, R.C. Atkinson, R.D. Luce, & P. Suppes (Eds.), Contemporary developments in mathematical psychology. (Vol. 2.) San Francisco: W.H. Freeman, 1974.

Ball, F., Wood, C., & Smith, E.E. When are semantic targets detected faster than visual or acoustic ones? Perception & Psychophysics, 1975, 7, 1-8.

Baron, J. Phonemic stage is not necessary for reading. Quarterly Journal of Experimental Psychology, 1973, 25, 241-246.

Baron, J., & Thurston, I. An analysis of the word-superiority effect. Cognitive Psychology, 1973, 4, 207-228.

Bjork, E.L., & Estes, W.K. Letter identification in relation to linguistic context and masking conditions. Memory & Cognition, 1973, 1, 217-223.

Blakemore, C. The baffled brain. In R.L. Gregory & E.H. Gombrich (Eds.) Illusion in nature and art. London: Duckworth, 1973.

Blesser, B., Shillman, R., Kuklinski, T., Cox, C., Eden, M., & Ventura, J. A theoretical approach for character recognition based on phenomenological attributes. International Journal of Man-Machine Studies, 1974, 6, 701-706.

Bouma, H. Visual recognition of isolated lower-case letters. Vision Research, 1971, 11, 459-474.

Breitmeyer, B.G., & Ganz, L. Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression and information processing. Psychological Review, 1976, 83, 1-36.

Broadbent, D.A. The hidden preattentive processes. American Psychologist, 1977, 32, 109-118.

Carr, T.C., Lehmkuhle, S.W., Kottas, B., Astor-Stetson, E.C., & Arnold, D. Target position and practice in the identification of letters in varying contexts: A word superiority effect. Perception & Psychophysics, 1976, 19, 412-416.

Cattell, J.M. The time it takes to think. Popular Science Monthly, 1888, 32, 488-491. Reprinted in James McKeen Cattell, Man of Science. Vol. 2, Addresses and formal papers. Lancaster, Pa.: The Science Press, 1947.

Chandler, J.P. Subroutine STEPIT finds local minima of a smooth function of several parameters. Behavioral Science, 1969, 14, 81-82.

Chomsky, N., & Halle, M. The sound pattern of English. New York: Harper & Row, 1968.

Cole, R.A. Listening for mispronunciations: A measure of what we hear during speech. Perception & Psychophysics, 1973, 13, 153-156.

Cole, R.A. & Jakimik, J. Understanding speech: How words are heard. Technical Report, Department of Psychology, Carnegie-Mellon University, 1977.

Cosky, M.J. The role of letter recognition in word recognition. Memory and Cognition, 1977, 4, 207-214.

Erickson, D., Mattingly, I.G. & Turvey, M.T. Phonetic coding of Kanji. Paper presented at the 83rd meeting of the Acoustical Society of America, Buffalo, New York, 1972.

Eriksen, C.W., Pollack, M.D., & Montague, W.E. Implicit speech: Mechanism in perceptual encoding. Journal of Experimental Psychology, 1970, 84, 502-507.

Estes, W.K. The locus of inferential and perceptual processes in letter identification. Journal of Experimental Psychology: General, 1975, 104, 122-145.

Estes, W.K. Bjork, E.L., & Skaar, E. Detection of single letters and letters in words with changing versus unchanging mask characters. Bulletin of the Psychonomic Society, 1974, 3, 201-203.

Foss, D.J., & Swinney, D.A. On the psychological reality of the phoneme: Perception, identification, and consciousness. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 245-257.

Frederiksen, J.R. Decoding skills and lexical retrieval. Paper presented at Psychonomic Society, St. Louis, November, 1976.

Frederiksen, J.R., & Kroll, J.F. Spelling and sound: Approaches to the internal lexicon. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 361-379.

Fujimura, O. Syllable as a unit of speech recognition. IEEE transactions on acoustics, speech, and signal processing, 1975, ASSP, 23, 82-87.

Gough, P.B. One second of reading. In J.F. Kavanagh & I.G. Mattingly (Eds.), Language by ear and by eye. Cambridge, Mass., MIT Press, 1972, p.

Gough, P.B., & Cosky, M.J. One second of reading again: In N.J. Castellan, Jr., D.P. Pisoni, and G.R. Potts (Eds.), Cognitive theory, Vol. 2. Hillsdale, N.J., Erlbaum, 1976.

Green, D.W., & Shallice, T. Direct visual access in reading for memory. Memory and Cognition, 1976, 4, 753-758.

Hawkins, H.L., Reicher, G.M., Rogers, M., & Peterson, L. Flexible coding in word recognition. Journal of Experimental Psychology: Human Perception and Performance, 1976, 2, 380-385.

Hubel, D.N., & Wiesel, T.N. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. Journal of Physiology, 1962, 160, 106-154.

Huey, E.B. The psychology and pedogogy of reading. New York: MacMillan, 1908, Republished by the MIT Press, Cambridge, Mass., 1968.

Jakobson, R., Rant, C.G.M., & Halle, M. Preliminaries to speech analysis: The distinctive features and their correlates. Cambridge, Mass.: MIT Press, 1961.

Johnson, N.F. On the function of letters in word identification: Some data and a preliminary model. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 17-29.

Juola, J.F., Choe, C.S., & Leavitt, D.D. A reanalysis of the word-superiority effect. Paper presented at the Psychonomic Society, Boston, Mass., November, 1974.

Keyser, S.J. & Halle, M. What we do when we speak. In P.A. Kolers and M. Eden (Eds.) Recognizing Patterns. Cambridge, Mass.: MIT Press, 1968.

Klapp, S.T. Syllable-dependent pronunciation latencies in number naming: A replication. Journal of Experimental Psychology, 1974, 102, 1138-1140.

Klapp, S.T., Anderson, W.G., & Berrian, R.W. Implicit speech in reading, reconsidered. Journal of Experimental Psychology, 1973, 100, 368-374.

Kleiman, G.M. Speech recoding in reading. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 323-340.

LaBerge, D. & Sammuels, S.J. Toward a theory of automatic information processing in reading. Cognitive Psychology, 1974, 6, 293-323.

Ladefoged, P. A course in phonetics. New York: Harcourt, Brace, and Jovanovich, 1975.

Lane, H., Boyes-Braem, P., & Bellugi, U. Preliminaries to a distinctive feature analysis of handshapes in American Sign Language. Cognitive Psychology, 1976, 8, 263-289.

Lesser, V.R., Fennell, R.D. Erman, L.D. & Reddy, D.J. Organization of the hearsay II speech understanding system. IEEE transactions on acoustic, speech, and signal processing, 1975, ASSP, 23, 11-23.

Liberman, A.M. Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. Perception of the Speech Code, 1976, 74, 431-461.

Lindsay, P.H., & Norman, D.A. Human information processing. New York: Academic Press, 1977.

Manelis, L. The effect of meaningfulness in tachistoscopic word perception. Perception & Psychophysics, 1974, 16, 182-192.

Marslen-Wilson, W. Linguistic structure and speech shadowing at very short latencies. Nature, 1973, 244, 522-523.

Marslen-Wilson, W.D. Sentence perception as an interactive parellel process. Science, 1975, 189, 226-228.

Marslen-Wilson, W., & Tyler, L.K. Processing structure of sentence perception. Nature, 1975, 257, 784-786.

Marslen-Wilson, W., & Welsh, A. Processing interactions and lexical access during word recognition in continuous speech. Unpublished manuscript, 1977.

Massaro, D.W. Perception of letters, words, and nonwords. Journal of Experimental Psychology, 1973, 100, 349-353.

Massaro, D.W. Understanding language: An information processing model of speech perception, reading, and psycholinguistics. New York: Academic Press, 1975. (a)

Massaro, D.W. Experimental psychology and information processing. Chicago: Rand-McNally, 1975. (b)

Massaro, D.W., & Cohen, M.M. The contribution of fundamental frequency and voice onset time to the /zi/-/si/ distinction. Journal of the Acoustical Society of America, 1976, 60, 704-717.

Massaro, D.W., & Klitzke, D. Letters are functional in word identification. Memory & Cognition, 1977, 5, 292-298.

Massaro, D.W., Venezky, R.L., & Taylor, G.A. Orthographic regularity, positional frequency, and visual processing of letter strings. Journal of Experimental Psychology: General, in press.

Mayzner, M.S., & Tresselt, M.E. Tables of single-letter and digram frequency counts for various word-length and letter-position combinations. Psychonomic Monograph Supplements, 1965, 1, 13-32.

McClelland, J.L. Preliminary letter identification in the perception of words and nonwords. Journal of Experimental Psychology: Human Perception & Performance, 1976, 2, 80-91.

McNeill, D., & Lindig, K. The perceptual reality of phonemes, syllables, words, and sentences. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 419-430.

Mezrich, J.J. The word superiority effect in brief visual displays: Elimination by vocalization. Perception & Psychophysics, 1973, 13, 45-48.

Nash-Weber, B. The role of semantics in automatic speech understanding. In D.G. Bobrow & A. Collins (Eds.) Representation and Understanding. New York: Academic Press, 1975.

Naus, M.J., & Shillman, R.J. Why a Y is not a V: A new look at the distinctive features of letters. Journal of Experimental Psychology: Human Perception & Performance, 1976, 2, 394-400.

Oden, G.C., & Massaro, D.W. Integration of place and voicing information in identifying synthetic stop-consonant syllables. WHIPP Report #1 Wisconsin Human Information Processing Program, July, 1977.

Pollatsek, A., & Carr, T. Wholistic and rule-governed encoding processes in word perception. Paper presented at Processing of Visible Language, Eindhoven, the Netherlands, 1977.

Reicher, G.M. Perceptual recognition as a function of meaningfulness of stimulus material. Journal of Experimental Psychology, 1969, 81, 275-280.

Rumelhart, D.E. Toward an interactive model of reading. In S. Dornic (Ed.) Attention and Performance VI, Hillsdale, N.J.: Erlbaum, 1977.

Savin, H.B., & Bever, T.G. The nonperceptual reality of the phoneme. Journal of Verbal Learning and Verbal Behavior, 1970, 9, 295-302.

Shillman, R.J., Cox, C., Kuklinski, T., Ventura, J., Blesser, B., & Eden, M. A bibliography in character recognition. Techniques for describing characters. Visible Language, 1974, 8, 151-166.

Sloboda, J.A. Decision times for word and letter search : A holistic word identification model examined. Journal of Verbal Learning and Verbal Behavior, 1976, 15, 93-101.

Sloboda, J.A. The locus of the word-priority effect in a target-detection task. Memory & Cognition, 1977, 5, 371-376.

Spector, A., & Purcell, D.G. The word superiority effect: A comparison between restricted and unrestricted alternative set. Perception & Psychophysics, 1977, 21, 323-328.

Spencer, H. The visible word. New York: Hastings, 1969.

Sternberg, S. The discovery of processing stages: Extensions of Donders' method. Acta Psychologica, 1969, 30, 276-315.

Taylor, G.A., Massaro, D.W., & Venezky, R.L. Visual information and orthographic structure in letter and word recognition. Paper presented at the Psychonomic meetings, Washington, D.C., 1977.

Theios, J., & Muise, J.G. The word identification process in reading. In N.J. Castellan, Jr., D.B. Pisoni, & G.R. Potts (Eds.), Cognitive Theory, Vol. 2, Hillsdale, N.J. Erlbaum, 1976.

Thompson, M.C., & Massaro, D.W. Visual information and redundancy in reading. Journal of Experimental Psychology, 1973, 98, 49-54.

Tweney, R.D., Herman, G.W., & Hoemann, H.W. Psychological processing of sign language: Effects of visual description on sign intelligibility. Journal of Experimental Psychology: General, 1977, 106, 255-268.

Venezky, R.L. Orthographic regularities in English words. Paper presented at Processing of Visible Language, Eindhoven, the Netherlands, 1977.

Venezky, R.L., & Massaro, D.W. The role of orthographic regularity in word recognition. In L. Resnick & P. Weaver (Eds.) Theory and practice of early reading. Hillsdale, N.J.: Erlbaum, in press.

Warren, R.M. Identification times for phonemic components of graded complexity and for spelling of speech. Perception & Psychophysics, 1971, 9, 345-349.

Wrolstad, M.E. A manifesto for visible language. Visible language 1976, 10, 5-40.

Zadeh, L.A. Quantitative fuzzy semantics. Information Sciences, 1971, 3, 159-176.